Co-aggregation Properties of trimeric autotransporter adhesins

Hawzeen Salah Khalil

Master thesis
Department of Biosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO
June 2018

Abstract

Trimeric autotransporter adhesins (TAAs) comprise a group of virulence-related proteins in Gram-negative bacteria. These obligate homotrimeric proteins are embedded in the outer membrane and function as adhesins. Members of this family bind to extracellular matrix components such as collagen and laminin and also confer serum resistance and autoaggregation. In order to investigate co-aggregation between different TAAs, we co-expressed a fluorescent label (sfGFP or mCherry) with a particular TAA and followed the interaction using fluorescent readout and microscopy.

We used two subtypes of TAAs: YadA from the enteropathogens *Yersinia enterocolitica* (YeYadA) and *Y. pseudotuberculosis* (YpYadA), and the immunoglobulin-binding Eib proteins from *Escherichia coli*, EibA, EibC, and EibD. The autoaggregation mediated by these proteins is homotypic (i.e. YadA binding to YadA, EibD binding to EibD etc.), but it is not known whether TAAs can mediate heterotypic interactions (e.g. YadA binding to EibA, i.e. co-aggregation between different TAAs.

Results show that there is co-aggregation between some populations expressing different TAAs, which can be explained by relatively high sequence similarity between the interacting TAAs in most cases, the level of co-aggregation correlated with the sequence similarity. However, in other cases, the TAAs did not interact despite high sequence similarity, showing exclusion of non-self-bacteria or the two different TAAs that did not co-aggregate.

We also performed biofilm assays for mixed population expressing different TAAs to see whether they form mixed biofilms or separate microdomains within the biofilm.

Our results showed that they are forming a mixed biofilm in most cases, but within exception of a few TAAs that showed segregation within the biofilm when mixed together.

In addition, we performed mutagenesis experiments to find out which residues or domain(s) in the TAAs are responsible for autoaggregation.

No individual point mutation or domain deletion abrogated the autoaggregation.

Acknowledgments

A master's thesis is a long journey. Therefore, this work and study, which came across many obstacles, would not have been possible without the help of those who first gave me the opportunity to do this study and who never stopped supporting and encouraging me, and those who taught me and answered my questions along the course of these work.

So, firstly, I would like to express my gratitude to Professor Dirk Linke, for accepting me into his group, and for being my co-supervisor.

Very special thanks to Doctor Jack C. Leo, my supervisor who helped me whenever I needed it, always showing great availability, creating a good working environment, teaching me and answering and clarifying all doubts that emerged throughout this work.

I want to thank Jonas Øgaard at the Research Institute for Internal Medicine, Oslo University Hospital, Rikshospitalet, for collaboration and making the script that I used for analysing the images taken by confocal scanning laser microscopy. I also thank Frode Skjedal at the IBV imaging facility for assistance with image acquisition and analysis.

I would, also, like to thank, everybody at the Linke lab and Leo group, who, likewise, when needed, answered my questions and helped me whenever possible.

Lastly, I would like to express my gratitude to my family especially my husband Hafez and my kids (Hemyar and Rozyar), and whole my family in Kurdistan for the support and encouragement.

Abbreviations

Ail	attachment and invasion locus
ATP	adenosine triphosphate
BAM	β-barrel assembly machinery
Cm	Chloramphenicol
CSLM	confocal scanning laser microscopy
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
ECM	extracellular matrix
Eibs	*Escherichia coli* immunoglobulin binding proteins
EspP	extracellular serine protease plasmid encoded
Fwd	Forward
Ig	immunoglobulin
IM	inner membrane
Kb	Kilo base
mCherry	monomeric Cherry fluorescent protein
nm	nanometer
OD_{600}	optical density at 600 nm
OM	Outer membrane
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PNAG	poly-N-acetylglucosamine
PORTA	Polypeptide transport associated
Rev	Reverse
RT	room temperature
SAAT	Self-associating autotransporters
SD	standard deviation
sfGFP	superfolder green fluorescent protein
SS	secretion system
ssDNA	Single- stranded DNA
STEC	shiga-toxin producing *E. coli*
TAA	trimeric autotransporter adhesin
UspA	ubiquitous surface protein A
VTEC	verotoxigenic *E. coli*
YadA	*Yersinia* adhesin A
YeYadA	*Yersinia enterocolitica* YadA
YLH	YadA-like head
YpYadA	*Yersinia pseudotuberculosis* YadA

Table of Contents

1 Introduction

1.1 Type V secretion systems (T5SS)

Gram-negative bacteria have a very different cell wall structure in comparison to Gram-positive bacteria. It consists of three layers: the innermost layer is named the inner membrane (IM), the middle layer which is the space between the IM and the outer membrane, called the periplasmic space or periplasm and contains a thin layer of peptidoglycan. The third, outermost layer is the outer membrane (OM).
In order to interact with the external environment, these bacteria possess a number of secretion systems to transport proteins to the cell surface or the extracellular medium. There are many different classes of secretion system in Gram-negative bacteria and in my study, I am focusing on the type V secretion system and more specifically on the type Vc subclass.

Type V secretion systems (T5SS) are divided into five subclasses, Type Va-Ve, as shown in (**Figure 1**). This scheme includes classical autotransporters (type Va), two-partner secretion systems (type Vb), trimeric autotransporter adhesins (type Vc), patatin-like autotransporters (type Vd), and inverse autotransporters (type Ve) (Fan *et al.*, 2016).
The simplest form of Type V secretion is known as the monomeric or classical autotransporter pathway (type Va secretion).
As the name implies, autotransporters (ATs) contain components that allow them to secrete themselves (Guérin *et al.*, 2017; Jain & Goldberg, 2007). Autotransporters contain three functional regions: a transmembrane β-barrel domain at the C-terminus that forms the outer membrane channel, a linker region between the passenger and the barrel, and a passenger that comprises the extracellular domain(s) of the protein and contains the functional part(s) of the autotransporter and may consist of several individual domains.

Sometimes, the passenger domain of some ATs is cleaved by an autoproteolytic reaction in order to release the passenger domain into the extracellular medium. So, an autoproteolytic reaction may happen in the barrel (e.g. extracellular serine protease, plasmid encoded (EspP)). But others are cleaved by exogenous proteases, e.g. NalP which cleaves other ATs in *Neisseria meningitides.* In contrast, AIDA-I does not possess a serine protease domain, but intramolecular cleavage of the passenger domain occurs through autoproteolysis that instead requires two acidic residues (Asp878 and Glu897) that reside in the passenger domain (Charbonneau *et al.*, 2009; Green & Mecsas, 2016; Leyton *et al.*, 2012; Velarde & Nataro, 2004).

Type V secretion systems are termed autotransporters because there is no known external energy source to drive the process of transport (I. R. Henderson *et al.*, 2004; Renn *et al.*, 2012). The only source of energy for transport available is the free energy of protein folding (Peterson *et al.*, 2010). But Kang'ethe & Bernstein suggested that the charge distribution is also the source for energy (Bernstein, 2015; Kang'ethe & Bernstein, 2013).

The type V secretion system is a two-step secretion system, in which proteins are first transported across the IM in an unfolded state by the Sec machinery.

Once they are in the periplasm, various chaperones such as Skp, SurA and DegP will surround the proteins and keep them in an unfolded state (Grijpstra *et al.*, 2013; Leo *et al.*, 2012; Rouviere & Gross, 1996; Schafer *et al.*, 1999; van Ulsen *et al.*, 2014).
The β-barrel is inserted into the OM via interaction with the help of the β-barrel assembly machinery (BAM) complex (Iadanza *et al.*, 2016; McCabe *et al.*, 2017; Noinaj *et al.*, 2013; Roman-Hernandez *et al.*, 2014; Schiffrin *et al.*, 2017).

Some type Va autotransporters are post-translationally modified; *e.g*. the AIDA-1 adhesin of *Escherichia coli* is glycosylated by a dedicated glycosyl transferase that is active in the cytoplasm (Benz & Schmidt, 1992). Another example is the NalP protease of *Neisseria meningitidis* which is lipid-modified during its transfer across the cell envelope (Pérez-Ortega *et al.*, 2017; van Ulsen *et al.*, 2014).

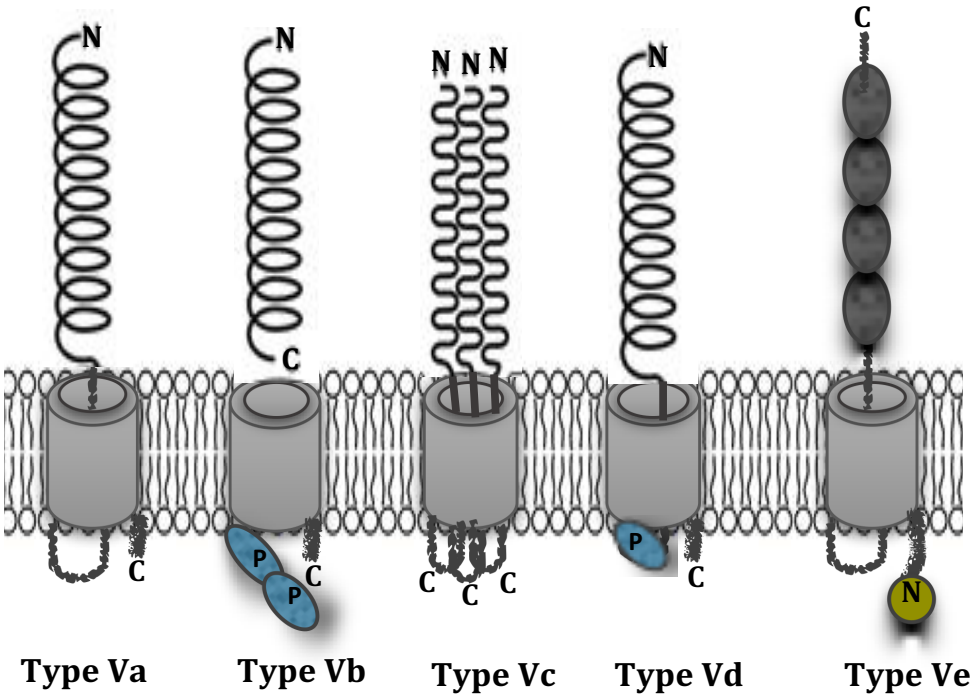


Figure 1. Type V secretion subtypes. Different subtypes of the Type V secretion system. The translocation domain is displayed in light grey consist of 12-stranded barrel for type V (a, e and c) and 16-stranded for type V (b and d), passenger domains in black and periplasmic domains in yellow. Periplasmic polypeptide transport-associated (PORTA) domain (small ovals in light blue) are labelled (P). The orientation of each protein is indicated by its N and C termini (denoted N and C). In type Va or monomeric autotransporters, the translocator domain and the passenger are expressed as a single polypeptide that also contains an N-terminal signal peptide. In contrast to classical autotransporters, the passenger and translocator functions in type Vb or TPSSs are located in separate polypeptide chains, but these are usually expressed from the same operon structure, and the β-barrel has two periplasmic domains called PORTA, which mediate protein –protein interaction (Clantin *et al.*, 2007). In trimeric autotransport (type Vc), passenger translocation is largely similar to classical autotransport, the major difference being the presence of three polypeptide chains rather than just one. Type Vd systems have an N-terminal passenger are expressed as a single polypeptide followed by one PORTA domain and a C-terminal passenger domain. Finally, in type Ve secretion or inverse autotransport, the extracellular C-terminal region is exported with the help of an N-terminal transmembrane β-barrel domain, which mediates binding to peptidoglycan (Leo *et al.*, 2015). This figure is made based on (Leo *et al.*, 2012).

1.2 Trimeric autotransporter adhesins

Trimeric autotransporter adhesins (TAAs) are obligate homotrimeric proteins and have type Vc secretion system. They are a widespread family of outer membrane proteins in Gram-negative bacteria. The polypeptides of trimeric autotransporters share a common molecular organization: each monomer contains an extended, highly variable N-terminal passenger and a conserved C-terminal translocation domain (Linke *et al.*, 2006).

TAAs follow a similar biogenesis pathway as classical autotransporters. In the first step of biogenesis, the signal peptide is recognized by the Sec machinery and mediates translocation through the IM. On entering the periplasm, various chaperones such as Skp, SurA and DegP will prevent the aggregation and folding of the TAA protein. The β-barrel the of the TAA is inserted in to the OM by interaction with the BAM complex.
Recently, Skidar *et al*. found that the β-barrel segments of TAAs fold into a trimeric structure in the periplasm. After a TAA is translocated into the periplasm, it will pass three steps: the first step is that three subunits rapidly form an asymmetric trimer in which two subunits fold into a structure that reflects their position in the final structure. Thereafter, this trimeric intermediate is targeted to the Bam complex and integrated into the OM. Finally, the third step is the initiation of passenger translocation, which is triggered by a relatively slow transition. After the passenger is rapidly translocated across the OM, the β-barrel forms into a heat-resistant and SDS-resistant structure (Sikdar *et al.*, 2017).

TAAs are important virulence factors in Gram-negative bacteria and acts as adhesins (Linke *et al.*, 2006). TAAs not only mediate adhesion to a variety of surfaces, but also mediate other virulence associated functions (**Table 1**).
As mediators of adhesion, they can bind to host cells, tissues, extracellular matrix (ECM) components, and also abiotic surface (Ishikawa *et al.*, 2012). In addition, some bind to molecules involved in immune responses, like immunoglobulins (Leo & Goldman, 2009; Sandt & Hill, 2001) factor H, and vitronectin (Biedzka-Sarek *et al.*, 2008; Capecchi *et al.*, 2005; Malito *et al.*, 2014; Muhlenkamp *et al.*, 2017).
YadA, a TAA found in both *Yersinia enterocolitica* (YeYadA) and *Y. pseudotuberculosis* (YpYadA), mediates binding to epithelial cells, macrophages and neutrophils (El Tahir & Skurnik, 2001), and has also an extensive ability to bind to ECM components such as fibronectin (Heise & Dersch, 2006; Tertti *et al.*, 1992) and collagen (El Tahir & Skurnik, 2001; Emody *et al.*, 1989; Heise & Dersch, 2006; Leo *et al.*, 2010; Nummelin *et al.*, 2004) .
BadA from *Bartonella henselae* mediates adherence to ECM and endothelial cells (Kaiser *et al.*, 2008; Muller *et al.*, 2011).
Ubiquitous surface proteins A (UspA) from *Moraxella catarrhalis*, UspA1 and UspA2, and Usp2H each possess a different function: UspA1 binds to carcinoembryonic antigen-related cell adhesion molecule 1 (CAECAM-1) (Conners *et al.*, 2008; Dje N'Guessan *et al.*, 2007), while Usp2 and Usp2H binds to ECM components such as collagen types I, II and III (Singh *et al.*, 2016).
Apa form *Actinobacillus pleuropneumoniae* mediates adherence to epithelial cells, specifically through the BD3 domain in the passenger (Xiao *et al.*, 2012).

Table 1. Examples of autoaggregating TAAs.

Organism	Protein	Functions (others than autoaggregation)	References
Actinobacillus pleuropneum	Apa	Adherence to host cells	(Cotter *et al.*; Xiao *et al.*, 2012)
Acinetobacter baumannii	AtaA	Autoagglutination Biofilm formation on (Biotic and abiotic surface)	(Ishikawa *et al.*, 2012)
Aggregatibacter actinomycetemcomitans	EmaA	Collagen binding	(Mintz, 2004)
Avibacterium	HMTp210	Haemagglutinationn, biofilm formation	(Wang *et al.*, 2014)
Bartonella henselae	BadA	Adhesion to host cells, binding to extracellular matrix proteins	(Kaiser *et al.*, 2008)
Bartonella quintana	VompA	Adhesion	(MacKichan *et al.*, 2008; Zhang *et al.*, 2004)
Escherichia coli	EibC EibC EibD EibG EibF EibA EibE	IgA and IgG binding, biofilm formation	(Leo *et al.*, 2011; Lu *et al.*, 2006; Sandt & Hill, 2001)
Escherichia coli		IgG binding, biofilm formation	(Leo *et al.*, 2011; Lu *et al.*, 2006; Sandt & Hill, 2001)
Escherichia coli	SAAT	Binding to epithelial cells	(Klemm *et al.*, 2006)
Escherichia coli	UpaG	ECM binding, biofilm formation.	(Valle *et al.*, 2008)
Haemophilus cryptic	Cha	Adherence to the maternal genital tract, and the neonatal respiratory tract.	(Sheets & St. Geme, 2011; Thanassi, 2011)
Haemophilus influenza	HadA	ECM binding, binding to and invasion of epithelial cells	(Serruto *et al.*, 2009)
Moraxella catarrhlis	MID(Hag)	IgD binding protein	(Pearson *et al.*, 2002; Riesbeck *et al.*, 2006)
Pasteurella(Pneumotropica	YadA like Protein (YadA_300)	Adherence and collagen binding	(Sasaki *et al.*, 2016)
Salmonella enterica	SadA	Epithelial cell binding Biofilm formation	(Grin *et al.*, 2014; Raghunathan *et al.*, 2011)
Veillonella atypica OK5	Hag1	Biofilm formation	(Peng Zhou *et al.*, 2015)
Yersinia enterocolitica	YeYadA	ECM binding. Serum and Phagocytosis resistance. Binding to epithelial cells	(El Tahir & Skurnik, 2001)
Yersinia pseudotuberculosis	YpYadA	ECM binding. Serum and Phagocytosis resistance. Binding to epithelial cells.	(El Tahir & Skurnik, 2001)

1.3 Model TAA proteins

The model protein used in this study were two groups of TAA, YadA from the enteropathogenic *Yersiniae*, YeYadA and YpYadA, and the immunoglobulin-binding Eib proteins from *Escherichia coli*, EibA, EibC, and EibD. All TAAs have similar structure, a lollipop-like shaped projections, on the bacterial surface (Hoiczyk *et al.*, 2000).
The structure consists of globular YadA like head domain, neck, stalk domain and conserved C-terminal domain. The structure of the YadA head and neck domain was one of the first structures representing a TAA head domain, solved by (Nummelin *et al.*, 2004). The structure of the EibD head and stalk was solved by (Leo *et al.*, 2011). Both TAAs have similar model structure (**Figure 2**) except for the N-terminal which is present in EibD and not in YeYadA, and there is no structure information available. Both TAAs have different functions, for example, YeYadA binds to collagen and the triple–helical conformation of collagen is required for binding but a specific sequence in collagen is not needed (Leo *et al.*, 2010; Leo *et al.*, 2008), while EibD does not bind to collagen, but binds to immunoglobulin (Ig) G.

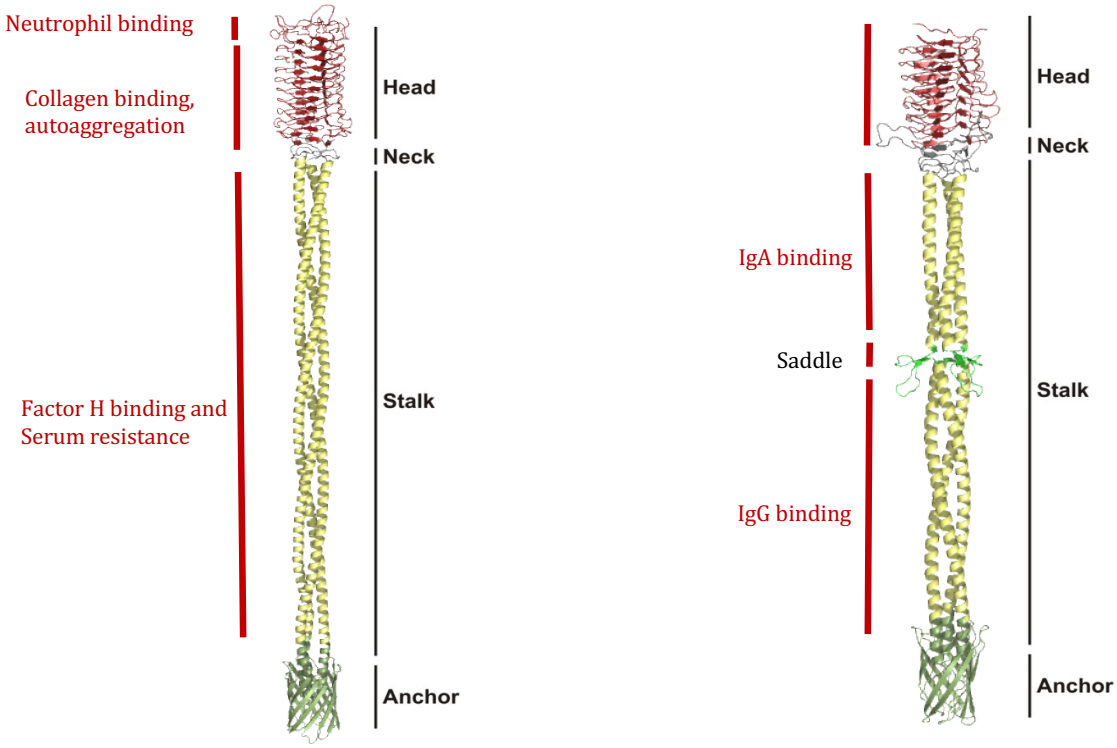


Figure 2. Computational models of TAA structures
YeYadA fiber is shown in the (**left**) (Koretke *et al.*, 2006); the figure was made using PyMol (Schrödinger). EibD fiber is shown in the (**right**), it has also an N-terminal, but there is no structure information available (Leo *et al.*, 2011). Both TAA have a similar model structure but different functions. The functional region is highlighted in red, while the structural region is highlighted in black.

1.3.1 The *Yersinia* YadA adhesins

Several species from the *Yersinia* genus are considered pathogens: *Y. enterocolitica* and *Y. pseudotuberculosis* are food-borne pathogens and causative agents of gastrointestinal infections, while *Y. pestis* is the agent of plague, a zoonotic disease that mainly affects rodents (Chain *et al.*, 2004; Cover & Aber 1989; Laporte *et al.*, 2015).
YadA forms rigid fibrous structures, which protrude approximately 23 nm from the cell surface (Hoiczyk *et al.*, 2000), and mediates adhesion to ECM (Leo *et al.*, 2012)

The *yadA* gene encodes the YadA protein and is carried on the pYV virulence plasmid. It is expressed by *Y. enterocolitica* and *Y. pseudotuberculosis* but not in *Y. pestis* which is not expressed at all and the reason for that is a frame shift in the *yadA* gene (Bolin *et al.*, 1982; El Tahir & Skurnik, 2001; Linke *et al.*, 2006; Reuter *et al.*, 2014; Skurnik & Wolf-Watz, 1989). The expression of the *yadA* gene is regulated and induced by a temperature of 37^0 C (El Tahir & Skurnik, 2001; Skurnik & Toivanen, 1992). Many functions, such as serum resistance, autoaggregation, phagocytosis resistance, invasion has been attributed to this adhesin (Balligand *et al.*, 1985; Skurnik *et al.*, 1984; Tertti *et al.*, 1992).

For YeYadA, this adhesin has important function, it binds to various types of fibrillar collagen, including types I, II, III, V, and also the network forming collagen type IV (Leo *et al.*, 2008; Schulze-Koops *et al.*, 1992). In contrast, YpYadA binds to fibronectin and laminin instead of collagen (Heise & Dersch, 2006). This is due to a 31 amino-acid (position 53 to 83) extension in the head domain called the uptake region.
Furthermore, the YeYadA also mediates adhesion to different types of cells, such as epithelial cells, macrophages, and neutrophils (Leo & Skurnik, 2011). It has the ability to block the three pathways that activate the complement system (the classical, lectin, and alternative pathways) that lead to opsonisation and lysis of bacteria (Biedzka-Sarek *et al.*, 2008; Lambris *et al.*, 2008; Mühlenkamp *et al.*, 2015). The ability of YadA to bind to collagen is crucial to the virulence of Ye, as its absence causes the bacteria to be avirulent in a mouse model. However, YadA is not an essential virulence factor for Yp (Pepe *et al.*, 1995; Roggenkamp *et al.*, 1995).

1.3.2 The immunoglobulin binding protein (Eibs)

Escherichia coli Ig-binding proteins (Eibs) were identified first in commensal *E. coli* strains by the ability bind soluble antibodies in a non-immune manner, which means that the mechanism does not require antibody-antigen interaction (Sandt *et al.*, 1997).
Until now, there are seven types of Eibs protein: EibA, C, D, E, F, G, and H described. The first four genes, *eibACD* and *eibE* were found in the *E. coli* strain ECOR9, and the eibF gene in E. coli strain ECOR2 (Sandt & Hill, 2000, 2001). In contrast, the gene encoding EibG was found in Shiga-toxin producing E. coli (STEC) serogroup O91 (Lu *et al.*, 2006). Later (Merkel *et al.*, 2010), found that EibG was also expressed by a number of other STEC strains of multiple serotypes which were lacking the gene encoding intimin.
And finally, the *eibH* gene was found from verotoxigenic *E. coli* (VTEC) which showed 88% identity with the *eibG* gene (Bardiau *et al.*, 2010).

All the Eibs binds to human IgG via the Fc region of the antibody. They also bind to IgA, except for EibA and EibE. However, none of Eibs protein showed any kind of binding to IgE and IgM (Sandt & Hill, 2000, 2001). Currently, there are no data on the Ig-binding abilities of EibH.
In addition to Ig-binding activity, all Eib proteins have another function. They act as adhesins: EibG causes a "chain-like adhesion" (CLA) phenotype when adhering to mammalian epithelial cells, a property specific to EibG and not seen in other types of Eibs (Lu *et al.*, 2006). EibG have three different subtypes: EibG (α, β, and γ), and they differ from each other in length of chain-like phenotype and adherence: EibG-α and EibG-β, respectively, displayed a typical chain-like adherence pattern (CLAP), by forming a long chain on both human and bovine intestinal epithelial cells. While, strains with EibG- γ adhered in short chains, a pattern which termed atypical CLAP (Merkel *et al.*, 2010).
Eibs proteins mediate serum resistance by an unknown mechanism (Lu *et al.*, 2006; Sandt & Hill, 2001). Lastly, EibG and EibD mediate autoaggregation, and EibD promotes biofilm formation (Leo *et al.*, 2011; Lu *et al.*, 2006).

1.4 Bacterial aggregation

Many bacteria, both environmental and pathogenic, have the property of autoaggregation (Trunk *et al.*, 2018). It is a distinct phenotype that can be visualized both macroscopically as flocculation and settling of bacteria cells in static conditions, and microscopically as aggregates or clumps of bacteria (**Figure 3**). There are two types of aggregation: autoaggregation and co-aggregation, as described below. -

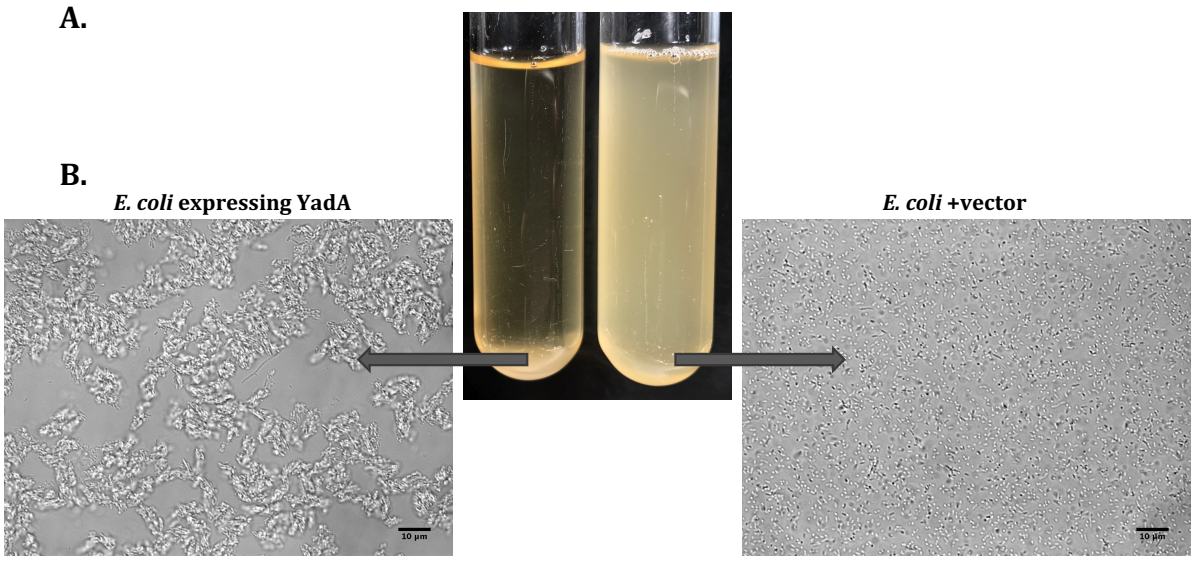


Figure 3. Bacterial autoaggregation (**A**) Macroscopic analysis of autoaggregation. *E. coli* cells expressing YadA (left tube) aggregate and settle at the bottom of the culture tube under static incubation, whereas an empty vector control culture (right tube) remains turbid. (**B**) Microscopic analysis of autoaggregation using phase contrast microscopy. Control cells (right micrograph) remain single, whereas YadA-expressing bacteria clump and form tightly packed aggregates (left micrograph). Scale bar=10µm. Based on (Trunk *et al.*, 2018).

1.4.1 Autoaggregation

Autoaggregation is the spontaneous clumping of bacteria which belong to the same strain (Schembri *et al.*, 2001), and it facilitates communication and biofilm formation in Gram-negative bacteria, both phenomena involved pathogenesis pathways (Zhang *et al.*, 2004). Cell-cell interactions allow the bacteria to form aggregates (Kjærgaard *et al.*, 2000; Travier *et al.*, 2013; Zhou *et al.*, 2016). To analyse auto-aggregation, the sedimentation of bacterial suspensions can be measured, as bacterial aggregates tend to settle (Trunk *et al.*, 2018).

There are many factors that mediate autoaggregation, including physicochemical characteristics of the cell surface such as hydrophobicity may affect autoaggregation and adhesion of bacteria to different surfaces (Kos *et al.*, 2003). In addition, surface factors of bacteria can mediate autoaggregation, these are termed autoagglutinins (Trunk *et al.*, 2018). Typical autoagglutinins are surface proteins, like TAA proteins (**Table 1**), but also other macromolecules can act as autoagglutinins: carbohydrates can mediate autoaggregation, such as the exopolysaccharide poly-N-acetylglucosamine (PNAG) of staphylococci (Formosa-Dague *et al.*, 2016). Another example is from *Campylobacter jejuni*, where the autoaggregative phenotype is dependent on glycosylation of flagella (Guerry, 2007).

In addition to TAAs, also non-TAA proteins can mediate autoaggregation (**Table 2**). Examples include Antigen 43 (Schembri & Klemm, 2001; Ulett *et al.*, 2006) and FimH (Schembri *et al.*, 2001), both from *E. coli*.

Almost all TAAs can tightly adhere to matrix components and host cells under both static and dynamic flow conditions (Muller *et al.*, 2011) and furthermore, this attachment can withstand high forces (El-Kirat-Chatel *et al.*, 2013), measured adhesion forces for autoaggregation using of a TAA from *Burkholderia cenocepacia* strain K56-2, and found that this adhesin forms homophilic *trans* interactions engaged in bacterial aggregation.

Table 2. List of selected characterized non-TAA proteins that have the autoaggregation properties based on (Trunk *et al.*, 2018).

Organism(s)	Protein	Class of protein	References
Aggregatibacter actinomycetemcomitans	FIp	Type IV pilus	(Henderson *et al.*, 2010)
Escherichia coli	TibA	Self-association autotransporter (SAAT)	(Sherlock *et al.*, 2005)
	Antigen 43	SAAT	(Heras *et al.*, 2014; Kjærgaard *et al.*, 2000; Schembri & Klemm, 2001; Ulett *et al.*, 2006)
	AIDA-1	SAAT	(Sherlock *et al.*, 2004)
	FimH	Type 1 fimbria, D-mannose specific adhesin	(Klemm & Schembri, 2000; Schembri *et al.*, 2001; Schembri & Klemm, 2001)
	Hra1	β- barrel protein	(Glaubman *et al.*, 2016)

Heamophilus influenzae	Hap	SAAT	(L. *et al.*, 2003)
Lactobacillus plantarum	D1	LysM-containing serine/therionine-rich protein	(Hevia *et al.*, 2013)
Legionella pneumophila	LcI	Collagen-like protein	(Abdel-Nour *et al.*, 2014)
Myxococcus xanthus	Pil	Type IV pilus	(Wu *et al.*, 1997)
Neisseria gonorrhoeae	Pil	Type IV pilus	(Park *et al.*, 2001)
Neisseria meningitidis	Aut A Pil	SAAT Type IV pilus	(Arenas *et al.*, 2015; Pérez-Ortega *et al.*, 2017)
Pseudomonas aeruginosa	PAK	Type IV pilus	(O'Toole & Kolter, 1998)
Rhizobium leguminosarum	RapA1	Rap family protein	(Ausmees *et al.*, 2001)
Salmonella enterica	SE17	Curli	(Collinson *et al.*, 1993)
Sinorizobium meliloti	EPSII	Exopolysaccharide	(Sorroche *et al.*, 2012)
Staphylococcus aureus	PNAG SasG	Exopolysaccharide MSCRAMM	(Formosa-Dague *et al.*, 2016; Kuroda *et al.*, 2008)
Staphylococcus epidermidis	Aap	MSCRAMM	(Rohde *et al.*, 2005)
Streptococcus pyogenes	M1	M protein	(Frick *et al.*, 2000)
Vibrio cholera	TCP	Type IV pilus	(Chiang *et al.*, 1995)
Xanthomonas campestris	FimA	Type IV pilus	(Ojanen-Reuhs *et al.*, 1997)
Yersinia pestis	YapC YPO0502 Ail (OmpX) attachment and invasion locus	SAAT HCP OmpX family β- barrel	(Felek *et al.*, 2008; Kolodziejek *et al.*, 2010; Podladchikova *et al.*, 2012)

1.4.2 Bacterial co-aggregation

Co-aggregation is the specific recognition and adhesion of genetically distinct bacteria. Specificity is mediated by complementary protein or polysaccharide agglutinins on the cell surface of aggregating cells (Kolenbrander *et al.*, 2002; Rickard *et al.*, 2004; Rickard, McBain*, et al.*, 2003).

This phenomenon is different from autoaggregation, which is the recognition and adhesion of genetically identical bacteria or (genetically very similar bacteria) (Rickard *et al.*, 2004; Rickard, Gilbert*, et al.*, 2003; Van Houdt & Michiels, 2005). Gibbons and Nygaard were the first demonstrated the co-aggregation between human dental plaque bacteria (Gibbons & Nygaard, 1970). The ability of bacterial cells to recognize and communicate with one other, leading to co-aggregation, is extensively investigated with regard to oral biofilms. A large amount of literature exists on the types and mechanisms

of interactions in bacterial tooth plaque (Elliott *et al.*, 2006; Kolenbrander *et al.*, 1985): There are also some reports on co-aggregation of organisms in the urogenital tract (Malik *et al.*, 2003), and it has also been shown that co-aggregation occurs between bacteria isolated from the human intestinal tract (Kos *et al.*, 2003; Reid *et al.*, 1988).
Relatively few studies of co-aggregation between aquatic biofilm bacteria (freshwater biofilms) and wastewater flocs have been reported (Rickard, McBain*, et al.*, 2003; Simoes *et al.*, 2008). Environmental factors such as substrate gradients, chemical or physical stress, and predation are known to trigger bacterial aggregation (Buswell *et al.*, 1997; Klebensberger *et al.*, 2006).

Among TAAs, a gene from the Gram-negative coccoid bacteria *Veillonela atypica hag1*, which encodes a YadA-like TAA, is involved in co-aggregation with the initial dental colonizers *Streptococcus gordonii, Streptococcus oralis* and *Streptococcus cristatus*, and the periodontal pathogen *Porphyromonas gingivalis*. The *hag1* mutant also abolished adherence to human buccal cells when the adherent bacteria were subjected to various chemical or physiological treatments, which suggest different mechanisms being involved in co-aggregation with different partners. The Hag1 proteins consist of 7178 aa and making it the largest bacterial surface protein reported thus far (Peng Zhou *et al.*, 2015).

1.5 Biofilm formation

In general, the key event in bacterial pathogenesis on the host tissue is adherence and colonization (Sherlock *et al.*, 2004). Biofilm is a surface-attached community of bacterial cells embedded in a self-produced polymeric matrix (Wolska *et al.*, 2016). These microbial collectives are found to be ubiquitous in almost every environment (Parsek & Singh, 2003).
Biofilms can be present on liquid surfaces as a floating mat, or submerged in the medium (Gupta *et al.*, 2016) and also on the surface of medical devices (Donlan, 2001).
The thickness of bacterial biofilm can vary from a single layer to multiple layers in which bacteria are attached to both the surface and to adjacent bacteria by an extracellular matrix consisting of polysaccharides, protein, and extracellular DNA (Hall-Stoodley & Stoodley, 2009; Karatan & Watnick, 2009; Satpathy *et al.*, 2016).

The formation of biofilms, in general, occurs when bacteria switch from a planktonic (free-swimming) state to a surface-attached state, and it occurs in multiple stages starting from the initial attachment followed by microcolony and macrocolony formation. Attachment of bacterial cells to abiotic surfaces and aggregation into microcolonies are considered the first step of biofilm formation, and cell surface hydrophobicity and motility play important roles in bacterial attachment (Li *et al.*, 2017; Stoodley *et al.*, 2002; Tribedi & Sil, 2014). Followed by microcolony and macrocolony formation, the final stage is the detachment by which bacteria return to the planktonic state again (Donlan, 2001; Gupta *et al.*, 2016).
Cell-cell interactions during biofilm production are crucial in determining biofilm architecture (Martínez-Gil *et al.*, 2010). These interactions are often mediated by adhesins located on the surface of the bacteria (Klemm & Schembri, 2000), which lead the bacteria to form microcolonies and biofilm in two ways (**Figure 4**).

In the first way, a single planktonic bacterial cell attaches to the substrate surface. The motility factors such as flagella or expression of surface adhesins play a role in the attachment of bacterial cell to the substrate (O'Toole & Kolter, 1998; Stoodley *et al.*, 2002). Alternatively, these single bacterial cells recruit other bacterial cells from suspension referred to as co-adhesion (Bos *et al.*, 1999). Just as soon, single cells can migrate along the substrate surface, e.g. using type IV pili, and aggregate (Dunne, 2002).

The second way that autoaggregation can initiate the biofilm is that the cells autoaggregate in the solution and then the aggregate settles on the surface (Kragh *et al.*, 2016). Both pathways lead to biofilm formation, and both may simultaneously play a role.

Aggregated cells have a competitive advantage over single cells at high cell densities. If the bacterial cell positioned at the top of the aggregate, they have more access to nutrients. However, the aggregated cells are at a disadvantage at low cell densities, because the cells in the middle of the aggregate have a limited nutrient access (Kragh *et al.*, 2016). The shape of the aggregate is also predicted to affect competition: at higher cell densities rounded aggregates fare better, but when the competition is low, spread aggregates that maximize surface area have an advantage (Melaugh *et al.*, 2016).

In addition to autoaggregation, some TAAs are known to promote biofilm formation. Biofilms can form on artificial surfaces like glass and plastic (Ishikawa *et al.*, 2012; Wolska *et al.*, 2016), but they can also form on biotic surfaces such as the accumulation of *Y. pseudotuberculosis* YPIII on the surface of *Caenorhabditis elegans* (Tan & Darby, 2004).

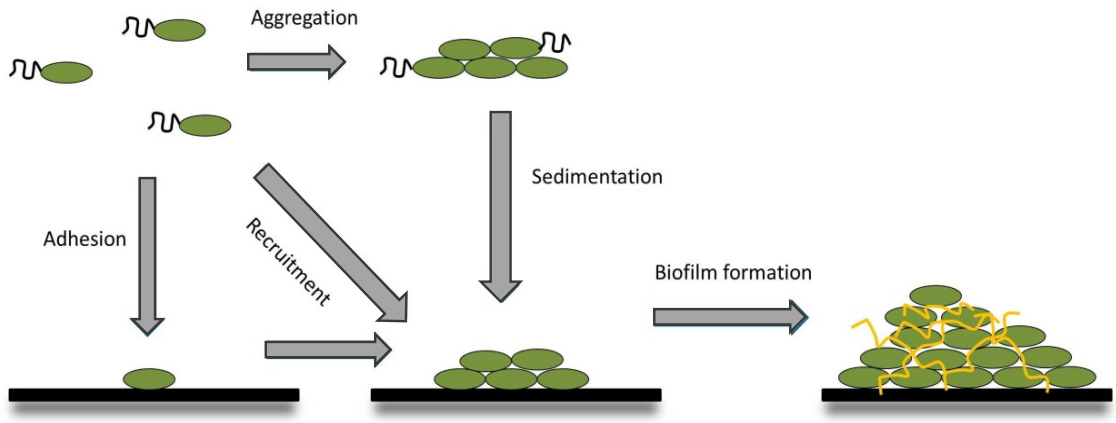


Figure 4. The role of autoaggregation in biofilm formation. Autoaggregation can lead to biofilm formation in two ways: planktonic bacteria can either attach to a substrate surface as single cells and then recruit more planktonic cells via aggregation to form a single microcolony, or planktonic cells aggregate in suspension and then settle on the substrate surface. Both pathways can lead to the formation of biofilm, Image reproduced from(Trunk *et al.*, 2018) under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0.

1.6 Aims of the project

The aims of this study were: -

1. To characterizes whether different TAAs co-aggregate with each other. It is not known whether TAAs co-aggregate or not. For this, we used two subtypes of TAAs: YadA from the *Yersiniae* (YeYadA) and (YpYadA), and the immunoglobulin-binding Eibs protein from *Escherichia coli*, EibA, EibC, and EibD.
Both YadA and the Eibs are known to mediate homotypic autoaggregation. (i.e. YadA binding to YadA, EibA binding to EibD etc.).

2. To find out whether populations expressing different TAAs form microdomains within the biofilm.

3. To find out which domains in the TAAs are responsible for autoaggregation.

2 Materials and methods

2.1 Genetics

2.1.1 Bacterial strains

Escherichia coli TOP10 (Invitrogen) was used for cloning and plasmid DNA amplification and storage. The expression strain *Escherichia coli* BL21(DE3) expresses T7 polymerase under the inducible lacUV5 promoter from (Novagen).

2.1.2 Primers

All the primers were designed manually and produced by Life Technologies, except for the primers used for amplifying both sites of the pACYCDuet-1 from Sigma -Aldrich®. The melting temperature (Tm) have been calculated by program OligoCalc (biotools.nubic.northwestern.edu/). The primer sequences used for amplification can be found in the (**Appendix 2, Table 1**).

2.1.3 Construction of a plasmid for co-expression

The plasmid used in this study for co-expression of TAAs and fluorescent markers was pACYC-Duet-1 from Novagen. The DNA templates used in this study for amplification of TAA coding sequences were from (Mikula *et al.*, 2012), except for YpYadA strain YPIII which was amplified from purified DNA. This plasmid contains two T7 RNA polymer promoter and two multiple cloning sites. To produce the plasmids in (**Table 3**), I cloned a TAA into one multiple cloning site and a fluorescent protein (mCherry or sfGFP) into the other. For selection, it has a chloramphenicol resistance gene.
All the constructs were made by using Gibson assembly which is a method where several DNA fragments with complementary overlaps can be cloned together with one step (Gibson *et al.*, 2009). The method requires linear PCR product for the insert and linearized plasmid as a vector, which in my case was pACYCDuet-1 (Novagen). Using the Gibson assembly master mix (**Appendix 4**) which contains a 5´ exonuclease (T5 exonuclease) which first, will create the 3´ single-stranded overhangs at the end of double stranded DNA of inserts and linear plasmid then, the overlapping ends anneal and the gaps formed by the exonuclease are filled in by the aid of Phusion DNA polymerase. Finally, both PCR products can be ligated by the aid of DNA ligase to form a circular DNA molecule.
In this study, first, the primers were made to include the overlapping ends for both the vector pACYCDuet-1 plasmid and insert which were the TAAs or fluorescent proteins in order to make the linearized PCR product.
Since this plasmid has two cloning site, we did cloning two times. The first time was for the first cloning site into which we cloned the DNA coding for a fluorescent protein (sfGFP or mCherry; **Table 3**) using primers pACYCDuetMCS1 Fwd and vector Rev with the primers for the fluorescent proteins.
For the second cloning site the gene for a TAA was inserted (**Table 3**) using primers pACYCDuetMCS2 Fwd and vector Rev and primers for the TAAs.
The Gibson assembly reaction was assembled on ice in a PCR tube.

The reactions contained 200 ng of insert DNA and 100 ng of a linearized vector(pACYCDeut-1), then an equal volume of 2x Gibson master mix (**Appendix 4**) was added to the reaction.

The reactions were incubated at 50°C for 30 min to 1 hr. After the reaction was completed, it was treated with 1 µl *DpnI* (NEB) was added to the PCR reaction in order to get rid of the methylated DNA template. Then the reactions were incubated for 1 h at 37 °C.

The reactions that were treated with *Dpn*1 were checked again by running in a 1% agarose gel. The PCR reactions were purified by using MiniElute® PCR purification Kit (250) form (QIAGEN). The DNA was transformed into chemically competent cells, *E. coli* TOP10 (**Section 2.1.6**).

Table 3. An overview of constructs and non-mutated plasmids used in this study. The genes (TAAs /fluorescent proteins) that were cloned into the 1st and 2nd multi cloning sites of plasmid.

Plasmid constructs	Insert in the first multi cloning site (MCS1)	Insert in the second multi cloning site (MCS2)	Ref.
pACYCDuet-1	-	-	Novagen
pACYCDuet-EibA/ sfGFP	sfGFP	EibA	This study
pACYCDuet -EibC /sfGFP		EibC	
pACYCDuet -YeYadA /sfGFP		YeYadA	
pACYCDuet-EibD/ mCherry	EibD	mCherry	This study
pACYCDuet-YpYadA/ mCherry	YpYadA		
pACYCDuet-YeYadA /mCherry(control)	mCherry	YeYadA	This study
pACYCDuet-EibD /sfGFP(control)	EibD	sfGFP	This study
pACYCDuet- sfGFP(control)	sfGFP	-	This study
pACYCDuet-mCherry(control)	mCherry	-	This study

2.1.4 Agarose Gel Electrophoresis

In order to verify the sizes of DNA products from cloning and colony PCR, 1% and 0.8% agarose gel were used. These were made by dissolving 1 g/0.8 g agarose powder (seaKem®LE Agarose, LONZA) in 100ml/80 of 1x TAE (Tris-Acetate-EDTA) buffer and heated by using the microwave until the solution became clear. The stain SYBR®safe from Thermo fisher was added when casting the gel. For 100ml gel, 10µl of SYBR®safe were added before pouring the agarose gel. Before loading the DNA samples on an agarose gel, 6x DNA sample buffer (Thermo Scientific) was added such that the final concentration was 1x. Then the DNA samples were loaded into the wells of the gel. Also, 4µl of 1kb DNA ladder (Thermo Scientific) was used as a marker. The agarose gel was run for 30 min at 75 V. DNA bands were visualized with a (Blue GelPic LED Box) from Nippon Genetics Imager. All solutions and buffers recipes can be found in (**Appendix 4**).

2.1.5 Polymerase Chain Reaction (PCR) for plasmid and insert amplification

For DNA amplification, Polymerase Chain Reaction (PCR) was used. It consists of cycles of different temperatures. At the first stage, which is denaturation, the DNA template will denature to create single-stranded DNA(ssDNA).
During the second stage, the complementary primers will anneal with the template. At the extension step, DNA polymerase will produce the desired fragment of DNA.
In order to perform the reactions, we needed dNTPs, Phusion polymerase and corresponding buffers which all were from New England Biolabs (NEB).
For PCR reactions, I followed the protocol from NEB:

Mix	**Volume**
-10mM dNTP	1 µl
-MilliQ water	37µl
-5X Phusion reaction buffer/5XPhusion reaction buffer	10 µl
-Forward primer (100 µM)	0.5 µl
-Reverse primer (100 µM)	0.5 µl
-Phusion polymerase	0.5 µl
-DNA template (50-250) ng	0.5 µl
The total volume of reaction	50 µl

Then the reactions were performed by transferring them into PCR machine (from Analytik Jena) with the following cycling parameter: -

1.Initial denaturation	98°C	30 sec	
2.Denaturation	98°C	10 sec	
3.Primer annealing	62°C	10 sec	
4.Elongation	72°C	2-3 min per kb (kilobase)	Return to step 2, with repetition 24times
5.Final extension	72°C	5 min	
6.pause	12°C	∞	

2.1.6 Transformation into *E. coli* chemically competent cells

In this study, two types of *E. coli* strains were made chemically competent: *E. coli* TOP10 for cloning and *E. coli* BL21(DE3) Gold for protein expression (**Section 2.2**).
Chemically competent cells were made by growing the bacterial cells to an optical density (OD_{600}) of (0.3-0.5). Bacterial cells were then collected by centrifugation (Beckman Coulter SX4400) for 10min at 4000 x *g*, resuspended in ¼ of the original volume of ice cold 0.1 M $CaCl_2$ (Merck) and incubated on ice for 30 min. Bacterial cells were pelleted

again but this time at 4°C at 4000 x *g* and resuspended in 1/25 of the original volume in ice cold 0.1 M $CaCl_2$ (Merck). An equal volume of ice cold 60% glycerol (VWR) was added to the cells. The cells could be stored at -80°C or used directly for transformation.
I followed a standard protocol for transformation of chemically competent cells:
First, an aliquot of competent cells was thawed on ice for a few minutes, then 100 µl of cells were used for the transformation procedure. 5 µl of completed Gibson reaction (**Section 2.1.3**), or 50 ng plasmid DNA was added to the competent cells, the suspension was mixed and the reaction was then left on ice for 30 min. The reaction was then subjected to a heat shock in a water bath at 42°C for 45 sec. After that, the reaction was directly put on ice for 2 min. Following this, the bacteria were plated on LB agar supplemented with 25 µg/ml chloramphenicol (Cm) and incubated overnight at 37°C in the incubator.
For successful transformation, 500 µl of SOC medium (see **Appendix 4** for media) was added to the cells after heat shock step and they were incubated on ice for 2min.
In order to recover the cells the mixture was incubated at 37 °C with agitation at 400 rpm for 1 h, then the cells were pelleted by a 2-min centrifugation at 13000 rpm/min. The pelleted cells were resuspended in 100 µl of medium remaining in the tube and plated on LB agar supplemented with 25 µg/ml Cm and incubated overnight at 37°C in an incubator.

2.1.7 Colony PCR

Colony PCR used to check whether the DNA insert was present or absent in the construct that has been transformed into *E. coli* TOP10 chemically competent cells.
First, the colony PCR reactions were made by making a mix which include the following:

Mix	**Volume**
10mM dNTP	0.4 µl
MilliQ water	17µl
10X Taq reaction buffer(NEB)	2 µl
Forward primer (100 µM)	0.2 µl
Reverse primer (100 µM)	0.2 µl
Taq DNA polymerase(NEB)	0.2µl
DNA template	Picked colony
The total volume of reaction	20 µl

Individual transformants first were picked up by a pipette tip, then inoculated in the colony PCR reaction and after that streaked on an LB agar plate containing chloramphenicol and numbered.

Then the reactions were transferred into a PCR machine from (Analytik Jena) with the cycling parameter below: -

1.Initial denaturation	94°C	30 sec	
2.Denaturation	94°C	10 sec	
3.Primer annealing	50°C	30 sec	
4.Elongation	72°C	2-3 min per kb (kilobase)	Return to step 2, with repetition 24times
5.Final extension	72°C	5 min	
6.pause	12°C	∞	

After the reactions were finished, the samples were run in a 1% agarose gel (**Section 2.1.4**) for small size DNA fragment (one insert+ plasmid) and 0.8% for big size DNA fragments (one inserts +plasmid).
The colonies that had given a product of the expected size of the insert, did the plasmid mini perp for it. For doing a plasmid miniprep, an overnight culture (o/n) has been made the day before and it was 10 ml or more because this plasmid has a low copy number from 10-12 per cell (Chang & Cohen, 1978; Rosano & Ceccarelli, 2014; Sathiamoorthy & Shin, 2012). The next day, the plasmids were purified using the Miniprep kit from Qiagen, after which they were sent for sequencing (**Section 2.1.8**).

2.1.8 Sequencing

To verify the constructs that were cloned by Gibson assembly into pACYCDeut-1, plasmid samples were sent to GATC Biotech.
The sequencing reaction included 5µl at 90 ng/µl of DNA for purified plasmid or 5µl at 80 ng for purified PCR product and sequencing primer (**Appendix 2, Table 2**) at a concentration of 5 µM the total volume of the reaction was 10 µl.
The Biotech lab provides the LightRun ™ sequencing service which is based on Sanger dideoxynucleotide sequencing (Sanger & Coulson, 1975). For analysing the sequencing results, the APE program was used (http//biologylabs.utah.edu/jorgensen/wayned/ape/).

2.1.9 One step site-directed plasmid mutagenesis

A series of point mutations in the head domain of YeYadA (see **Figure 17** in results) were made in order to make a mutation, using one-step site-directed plasmid mutagenesis (Liu & Naismith, 2008). Also, deletion of some selected domains was done. fluorescent protein on the other site was used as the template (**Table 4**). A pair of complementary primers (**Appendix 2, Table 3**), both including the appropriate nucleotide sequence to yield the desired amino acid, were used for PCR.
For PCR reactions, I followed the protocol below

Mix	**Volume**
-10mM dNTP	1 µl
-MilliQ water	37µl
-5X Phusion reaction buffer/5XPhusion reaction buffer	10 µl
-Forward primer (100 µM)	0.5 µl

-Reverse primer (100 µM)	0.5 µl
-Phusion polymerase	0.5 µl
-DNA template (50-250) ng	0.5 µl
The total volume of reaction	50 µl

The cycling parameters below used for amplifying the DNA.

1.Initial denaturation	98°C 30 sec	
2.Denaturation	98°C 15 sec	
3.Primer annealing	60°C 20 sec	
4.Elongation	72°C 7 min per kb (kilobase)	Return to step 2, with repetition 24times
5.Final extension	72°C 5 min	
6.pause	12°C ∞	

In a, round of PCR cycles, these primers anneal to the template DNA, replicating the plasmid DNA with the mutation the derived PCR product was digested using 1 µl *DpnI* (NEB) which has been added to the PCR reaction in order to get rid of the methylated DNA, then the reactions were incubated for 1 h at 37 °C.
The reactions that were treated with *Dpn*1 were checked again by running in a 1% agarose gel. The PCR reactions were purified by using MiniElute® PCR purification Kit (250) form (QIAGEN). The PCR used for transformation into competent *E. coli* TOP10 (**Section 2.1.6**). The sequences of the resulting plasmids were verified as correct by DNA sequencing (**Section 2.1.8**).

For protein expression, plasmids were transformed into *E. coli* BL21(DE3) Gold, and transformed cells plated on LB media supplemented with 25 µg/ml Cm (for more details see (**Section 2.1.6**).

Table 4. An overview of all mutated plasmids used in this study

Plasmid constructs	Ref.
pACYCDuet-YeYadA **I94A /** sfGFP	This study
Q124A	
E80A	
N166A	
A165D	
L110A	
E182A	
K68A	
pACYCDuet-YeYadA **Delta YLH** /sfGFP	This study
pACYCDuet-EibD **Delta_ N** /mCherry	This study
pACYCDuet-EibD **Delta_ YLH**) /mCherry	This study
pACYCDuet-YpYadA **Delta uptake** /mCherry	This study

2.2 Induction of protein production

All induction cultures in this study were grown in a warm room at 30 °C.

2.2.1 Induction using Isopropyl β-D-1thiogalactopyranoside (IPTG)

Isopropyl β-D-1thiogalactopyranoside (IPTG) is a lactose analogue which induces the expression of genes regulated by the lac promotor (in this case, T7 RNA polymerase). As mentioned in (**Section 2.1.6**), for expression I used *E. coli*. BL21(DE3), which contains T7 RNA polymerase gene under the *lac* promoter. The expression of the protein is controlled by the *lac* repressor.
The pACYCDuet-1 plasmid used for making the constructs in this study has only two T7 promoter. Upon IPTG induction, the T7 RNA polymerase will bind to the plasmid pACYCDuet-1 promoter and initiate mRNA production of the TAA along with the fluorescent protein (sfGFP or mCherry).
The IPTG concentration was optimized using several concentrations (0.1, 0.5, and 1 mM) and a negative control (un-induced culture) was also included, in both Luria-Bertani (LB) and (PA-0.5G defined or clear media) (Studier, 2005) (**Appendix 4** for media).
Bacterial cultures were grown to an OD_{600} of 0.5, then IPTG (G Biosciences ®) was added at various concentrations to the culture. After an optimized time 2.5 h of induction a sedimentation assay (**Section 2.3**) was performed after induction.

2.2.2 Autoinduction

Autoinduction medium ZYP-5052 (Studier, 2005) was prepared as described in (**Appendix 4**). Autoinduction medium was used for co-expressing the *lac* promoter in *E. coli* BL21(DE3). The 5052 solution contains carbon source for bacteria: glycerol, glucose, and lactose.
The bacteria utilize the glucose first as a carbon source, which represses the breakdown of lactose, and induction from the *lac* promotor. This will let the bacterial culture to grow to a higher density before protein expression is induced. Once the glucose level is decreased, then bacterial cells will utilize the glycerol and lactose as the carbon source and the expression from the *lac* promotor will be induced.

2.3 Bacterial sedimentation assay

2.3.1 Bacterial sedimentation assay for measuring autoaggregation

Sedimentation assays are a quantitative method for measuring the autoaggregation. The measurement is done by measuring the fluorescence of the cultures from the top of the tubes at given intervals. The reduction in fluorescence is then plotted as a function of time.
The assay was started as shown in (**Figure 5**) by first growing an overnight culture (5 ml LB medium with 25 µg/ml Cm at 30 °C) for each individual construct made in this study (**Section 2.1.3, Table 3**). The next day, the overnight cultures were diluted 1:20 in 20 ml fresh LB supplemented with 25µg/ml chloramphenicol and 0.2% (w/v) Glucose. The cultures were grown with shaking 200 rpm at 37 °C until the OD_{600} reached ~0.5.
After that, the samples were placed into 30 °C warm room for 30 min with shaking at 150 rpm/min. 0.5mM IPTG was added to induce protein production. After 2.5 h of induction,

the samples were transferred into narrow tubes and the tubes incubated statically without any agitation at room temperature(RT).

200 µl samples were taken from the very top of each tube, transferred into 1.5 ml Eppendorf tubes and then centrifuged at 13,000 rpm for 1 min. The pellet was resuspended in 200 µl of 1x PBS and then put it into a black 96 well polystyrene plate (Greiner bio-one) in order to measure the fluorescence in a Synergy ™ H1 plate reader (BioTek). The fluorescence measured at regular intervals over time using plate reader (this process was done every 30 min) and for 8 reading within three biological replicates. For samples containing sfGFP, the excitation and emission wavelengths used were 483 and 510nm, respectively. The gain was 110. For mCherry, the excitation and emission wavelengths used were 580 and 610 nm, respectively, with gain 60. The samples with mCherry were read first to prevent bleed through from the GFP measurement.

To estimate autoaggregation, the fluorescence measured at each time point was compared to the to the fluorescent intensity at time point zero and results expressed as a percentage. For the negative control, a construct that included only the fluorescent protein without any TAAs was used (**Section 2.1.3**). For data plotting, three biological replicates were used and results were presented as the mean with standard deviation. At the end of each experiment, some of the sedimented cells were taken for imaging using confocal scanning laser microscopy (**Section 2.4.2**).

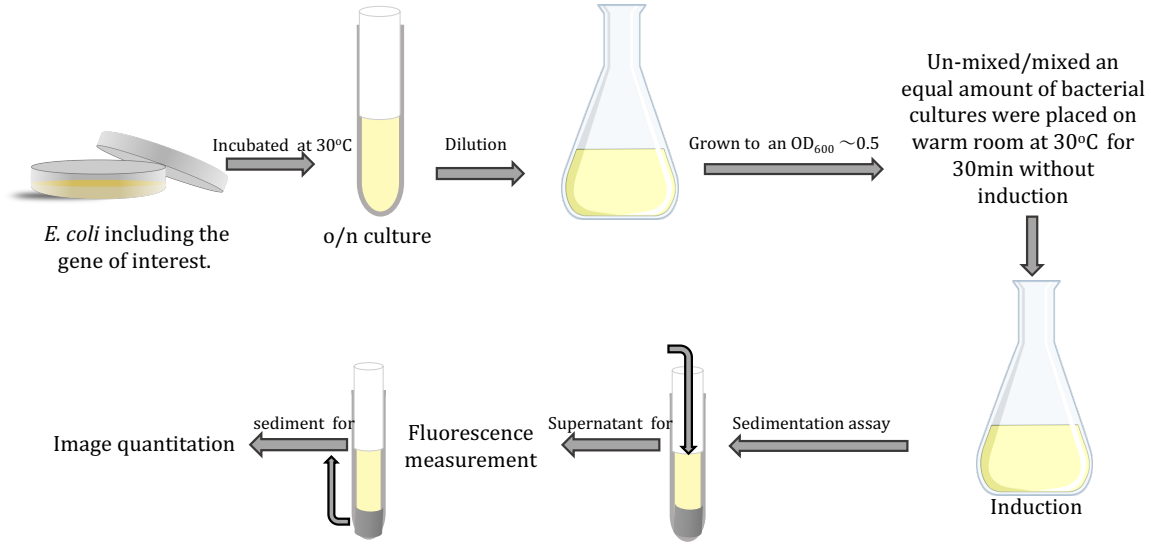


Figure 5. Simple diagram illustrating the experimental procedure for sedimentation assays.

2.3.2 Bacterial Sedimentation assay to measure co-aggregation

Quantitative analysis for measuring the co-aggregation was done the same way as in **Section 2.3.1**, but with some exceptions as shown in (**Figure 5**). After the bacterial culture reached the mid-log phase (OD_{600} of ~0.5), the OD_{600} of all samples was measured, and an equal amount of two bacterial cultures expressing different TAAs and fluorescent proteins were mixed together in a single flask (volume ~10 ml).

After measuring the OD_{600} for all samples the calculation below used in a final volume of 10 ml for mixing the bacterial culture together.
C_1= is the (red or green) bacterial culture with a higher OD_{600} concentration.
C_2= is the (red or green) bacterial culture with a lower OD_{600} concentration.
V_1=is the volume of culture 1(to be calculated).
V_2= is the volume of culture 2 (which is 5 ml).

$$C_1 \text{ x } V_1 = C_2 \text{ x } V_2$$

$$V_1 = \frac{C2 \text{ x } V2}{C1}$$

Thus, each culture expressing mCherry was mixed with one expressing sfGFP. After mixing, the mixed cultures were incubated at 30 °C for 30min with shaking at 150 rpm/min. The cultures were then induced by adding 0.5 mM IPTG to the ~10 ml of mixed bacterial culture. After 2.5 h of induction, the samples were transferred into narrow tubes and the tubes were incubated statically without any agitation at RT.
The rest of experimental procedure was the same as in (**Section 2.3.1**).
At the end of each experiment, some of the cell sediment was taken for imaging using CSLM (**Section 2.4.2**).

2.4 Microscopy

2.4.1 Phase contrast microscopy

In order to make sure that the bacterial constructs that were made in (**Section 2.1.3**) do not aggregate before inducing the bacterial culture with IPTG, I examined the bacteria by phase contrast microscopy. The images of the bacteria were taken by using the Axioplan 2 imaging microscope (ZEISS). Images were taken (**Figure 6**) at 63x magnification using a halogen lamp. Then images were further processed for display by using image J a free Java-based image-processing package (http://rsb.info.nih.gov/ij/).

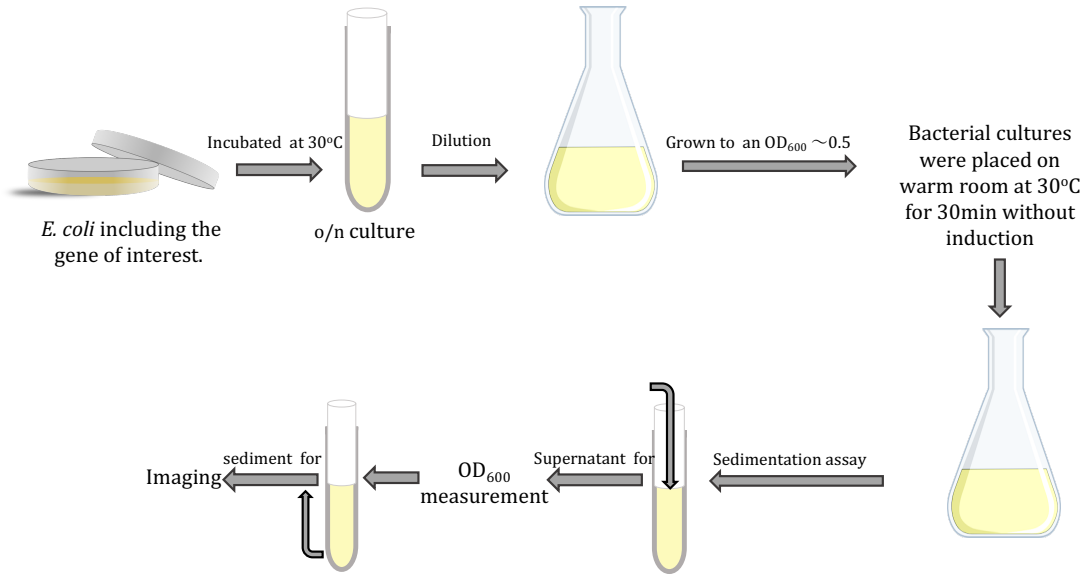


Figure 6. Simple illustration showing the experimental procedure for sedimentation assay for all un induced samples with IPTG for optimization process.

2.4.2 Confocal scanning laser microscopy (CSLM)

Directly after the sedimentation assays for both autoaggregation (**Section 2.3.1**) and co-aggregation (**Section 2.3.2**), some of the cell sediment was taken very carefully by using a long transfer pipette (catalogue nr. 612-2842 from VWR) and put on the top of a glass slide, then covered with a coverslip. The glass slide carrying the samples were inverted and put on the objective for examination. Microscopic observations and image acquisition were performed using an inverted confocal scanning laser microscope (Olympus Fluoview 1000) mounted with a PlanApo 60x/1.42 oil immersion objective (Olympus, Hamburg, Germany) and photomultiplier tube detectors. Bacterial cells were maintained in an incubator chamber while imaging that kept a stable environment with 37 °C, and 5% CO_2 levels. Flourechromes were excited with diode lasers (488, 559, and 635 nm).
Images were taken for each sample from six random fields representing at least two biological replicates. Images were further processed using ImageJ (**Section 2.5.1**).

2.4.3 Andor Dragonfly spinning disc confocal microscopy

Spinning disc confocal microscopy was used for it is speed for taking images of biofilms without destroying the structure of the biofilm and for obtaining high-resolution images. The 3D stacks were acquired on an Andor Dragonfly spinning disc confocal microscope equipped with an iXon 888 Ultra EMCCD camera and a Nikon Eclipse Ti inverted microscope. For this particular imaging, we used the Nikon PlanApo 60X 1.4 NA Oil immersion objective.
Typically, a 3D stack would consist of 350 frames in two colours sampled according to the Nyquist criterion for optimal 3D reconstruction. 3D biofilm images were taken for each chosen sample with three biological replicates.

2.5 Image analysis

2.5.1 Analysing 2D images taken by CSLM

2D images that were taken from co-aggregation samples by CSLM were analysed using a script made by Jonas Øgaard (Research Institute for Internal Medicine, Oslo University Hospital, Rikshospitalet). The script (**Appendix 3**), which was implemented in the image J analysis software (Abramoff *et al.*, 2003) analyses the images by segmentation through image J. After taking images by CSLM, the images were processed first by the image J software. The file for each image was converted from the oib format to a jpg file, then these jpg files were passed through the script. The script measures the fraction of green bacteria (expressing sfGFP) that are within 15 pixels of a red bacterium and vice versa. I refer to this percentage as the association index, which is given as a percentage of bacteria adjacent to a bacterium of the opposite colour compared to the total bacteria of that colour. The association index for each sample was calculated from six random fields representing at least two biological replicates.

2.5.2 Imaris XTension spot colocalization for analysing 3D biofilm images

Imaris form (Bitplane) software used to analyse the 3D images taken by the Andor Dragonfly confocal microscope. This software finds each green spot which represents a single bacterium fluorescing in the green channel that is colocalized with a red spot (bacterium with red fluorescence) within a distance of 1.15 µm. The distance based on the average of total width and length of the bacteria. The association index is given as a percentage of the total bacteria of one type calculated from three biological replicates.

The values have been found by using equation below, for example: -

$$\textbf{\% Red to green} = \frac{\text{number of colocalized red spots}}{\text{number of (colocalized red spots} + \text{ non}-\text{colocalized red spots)}} * 100$$

And the same equation was used for calculating the **% Green to red** association index**.**

2.6 Quantification of biofilms using crystal violet

Biofilm assays were performed using polystyrene 6-well plates (Sarstedt) and 96-well plate (Sarstedt). All the plasmids containing TAA genes that were made (**Table 3**), were transformed into *E. coli* Bl21(DE3) Gold. Cultures were grown in LB medium overnight, and the following day these were diluted in to 1:20 in 20 ml LB + 0.2% (w/v) Glucose + 25 µg/ml Cm and grown to an OD_{600} value of ~0.5 at 37 °C.
The OD_{600} was measured for each individual sample, then 1 µl of each bacterial culture along with 500 µl autoinduction medium (**Appendix 4** for media) supplemented with the 25 µg/ml Cm were added to the 6-well plates. Plates were incubated at 30 °C for 92 h,

either statically or with agitation at 40 rpm/min. Three biological replicates were made for each sample.
To measure biofilm formed on glass, coverslips were first placed inside the wells of a 6-well polystyrene plate, then the plate was incubated for 92h without agitation at 30 °C. The glass coverslips were moved to a new 6-well plate and then they were washed once with 500 µl 1xPBS. They were then stained with 0.5% (w/v) crystal violet by adding 500 µl of the solution to each well and incubating for 2 min. Then again the glass cover slips were moved to a new plate and were washed several times with 500 µl of 1xPBS until the control sample gave a clear colour. The stain from the biofilms was solubilized in 99% ethanol and after minutes, 200 µl of the solubilized dye for each individual sample were moved to 96-well plate (Sarstedt). The absorbance was read at 630nm using plate reader. The data were plotted as the mean with standard deviation for three replicates.
The staining was done the same way for biofilms formed on a polystyrene surface except for moving it to a new plate and also 96-well plate have been used.

2.7 Biofilm formation assay

The assay was started as shown in (**Figure 7**) by first growing an overnight culture (5 ml LB medium with 25 µg/ml Cm at 30 °C). For each individual construct were made in this study (**Section 2.1.3, Table 3**). The next day, the overnight cultures were diluted 1:20 in 20 ml fresh LB supplemented with 25 µg/ml chloramphenicol and 0.2% (w/v) Glucose. After the bacterial culture reached the log phase (OD_{600} of ~0.5). The OD_{600} of all samples was measured, and an equal amount of two bacterial cultures expressing different TAAs and fluorescent proteins were mixed together in a single flask (volume ~10 ml) (**Section 2.3.2**). Autoinduction medium (**Appendix 4** for media) was used for biofilm formation (**Sections 2.6**) and cultures were prepared by adding 3 µl of mixed bacterial culture with 3 ml of autoinduction media into 35 mm glass bottom culture plates coated with poly-D-lysine (MatTek).
These cultures plates were incubated for 92 h without agitation at 30 °C.
Images were taken for each sample using Andor dragonfly microscopy (**Section 2.4.3**).

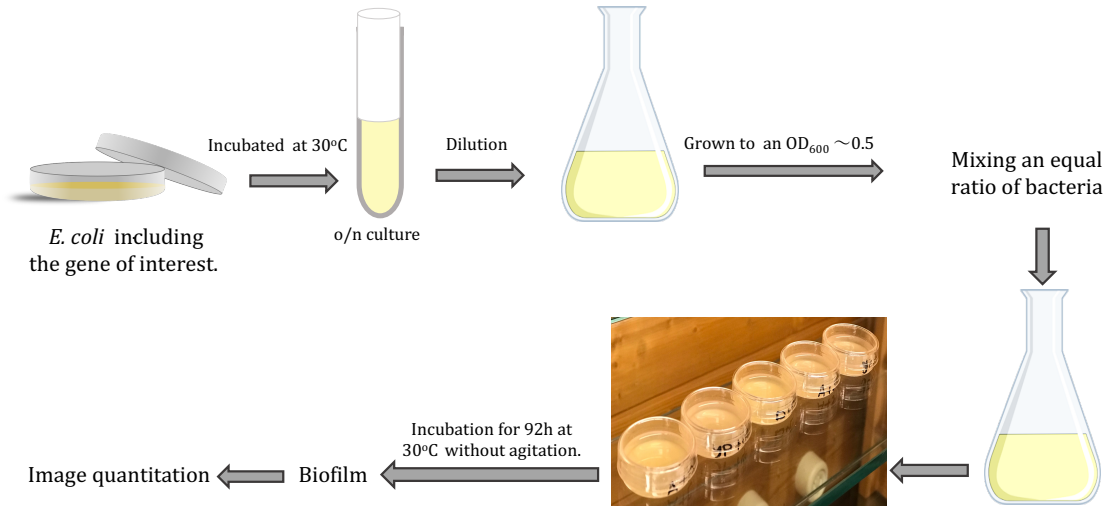


Figure 7. Simple illustration showing the experimental procedure for biofilm formation assay

2.8 Bioinformatics

Multiple sequence alignment is commonly used to analyse sets of homologous protein (Chatzou *et al.*, 2016). Clustal Ω (Alva *et al.*, 2016; Zimmermann *et al.*, 2017), part of the Tuebingen toolkit (https://toolkit.tuebingen.mpg.de/), was used to align multiple protein sequences (**Appendix 1**). In order to find the percentage of similarity between these different TAAs, we deleted the first 26 amino acid sequence for signal peptide and also the conserved membrane anchor was deleted. Then the aligned protein sequences were passed through another program that calculates the identity and similarity of each sequence pair (Stothard, 2000) (Ident and sim –bioinformatics .org).

2.9 Statistical analyses

All statistical data analysis was done using Microsoft Excel (Microsoft) by computing the reduction of fluorescence expressed as the percentage of initial fluorescence. The results were expressed as a mean with SD and plotted on the graph.
The graphs were made by using Graph Pad Prism version 7.0c (Graph Pad Software, La Jolla, Calif., USA).
The values in the results display the mean of N experiments calculated by using the equation below: -

$$\mu = \frac{\Sigma X}{N}$$

Where μ refers to the mean of the measurement, Σ is the summation (addition) sign, x is each individual value of measurement, and N is the number of experiments (Biological replicates).
The standard deviations (SD) displayed in the results were calculated using the equation below: -

$$SD = \sqrt{\frac{\Sigma(X-\mu)^2}{N-1}}$$

Where Σ is the summation (addition) sign, x is each individual value of measurement, μ refers to the mean of the measurement, and N is the number of experiment (Biological replicates).

3 Results

3.1 General strategy for investigating the co-aggregation of TAAs

The model TAAs used in this study were two subtypes of TAAs, YadA from the enteropathogens *Y. enterocolitica* (YeYadA) and *Y. pseudotuberculosis* (YpYadA), and the immunoglobulin-binding Eib proteins from *E. coli*, EibA, EibC, and EibD. Since the discovery of GFP by (Shimomura *et al.*, 2005), many research used the fluorescent marker in genetic studies related with following the pathway or function of the gene of interest (van Zyl *et al.*, 2015). In order to determine the co-aggregation properties of TAAs, the pACYCDuet-1 plasmid which has two multi cloning sites have been used and as mentioned in materials and methods (**Section 2.1.3**). A TAA was cloned in one site and a fluorescent protein (mCherry or sfGFP) on the other site. Once the bacteria co-expressed we will have two groups of bacteria that have a different colour. A red group and green group as in (**Figure 8**) below:

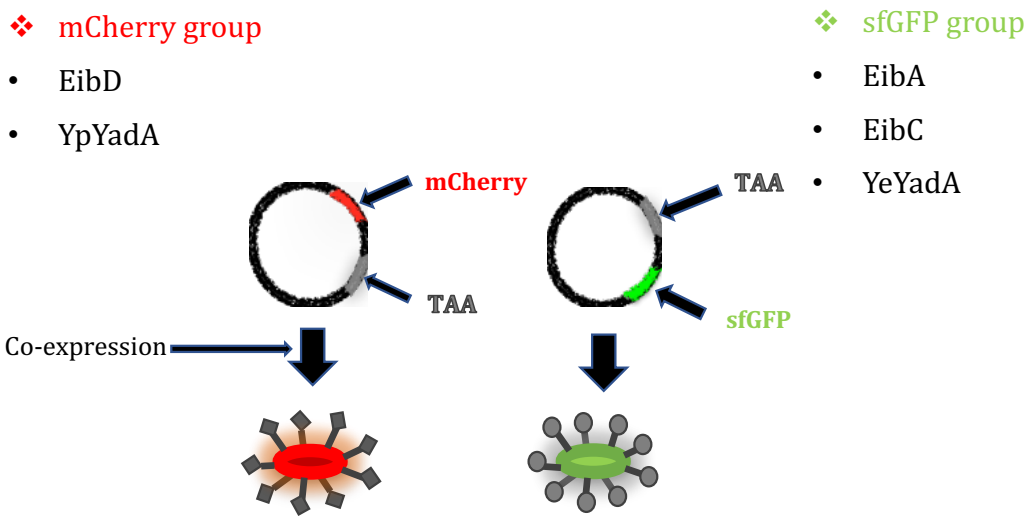


Figure 8. Simple illustration of the constructs using the pACYCDuet-1 plasmid

A negative control was also included which is a plasmid within a fluorescent protein without TAAs, and other controls were also included, which were the same type of TAA like EibD and YeYadA both combined with sfGFP and mCherry (**Section 2.1.3, Table 3** in material and methods).

In order to compare the translocation domain of different types of TAAs. A multiple sequence alignment for the TAA proteins used in this study was performed using the MPI Bioinformatics Toolkit (http://toolkit.tuebingen.mpg.de).

The purpose behind it is to know if the similarity of sequence has any effect on co-aggregation between the different mixed TAAs used in this study. Based on the protein sequence of each TAAs used in the study, the domain organization of the TAAs used in the study (**Figure 9**). The whole TAA molecules in general are lollipop-like structures consisting of a head, neck, stalk and conserved membrane anchor (Hoiczyk *et al.*, 2000).

The most widely studied TAA is YeYadA, and if we compare it to YpYadA we can see that the main difference is the uptake region, a sequence consisting of 31 amino acid residues which give this protein the preference for binding to fibronectin (Heise & Dersch, 2006) . All Eibs protein used in this study have a N-terminal region that follows the signal peptide, referred to as the Eib N-terminal region. This N-terminal region is not found in the YadA proteins and it has an unknown function and structure. All Eib proteins contain a saddle that is located half way through stalk domain. It is only 22 residues long and consists of three antiparallel β-sheets. is responsible for rotating the chain 120° clockwise (Leo *et al.*, 2011). In contrast to EibC and EibD, EibA lacks a YLH.

The values for similarity in (**Figure 9**) are only for the passenger (head, neck, and stalk. The results show that both YeYadA and YpYadA have a high percentage of sequence similarity (69%) and sequence identity (66%), and in the Eibs group, EibD and EibC share 89% sequence similarity and 83% sequence identity, while the sequence similarity between EibA and EibD is 37% and the sequence identity is 27%. The similarity between YeYadA and EibD is 25% and the sequence identity is 17%.

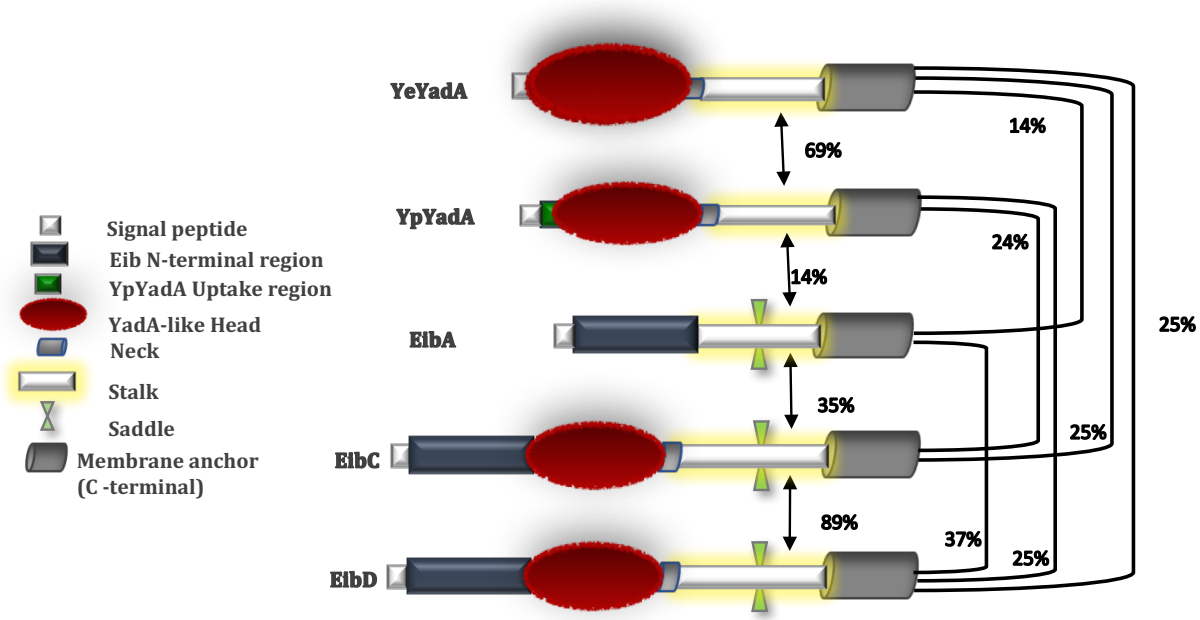


Figure 9. Schematic representation showing the relative sizes and domain organization of the model TAAs used in the study. In general, the whole TAA molecules are lollipop-like structures consisting of a head, neck, stalk and conserved membrane anchor. The head consist of a single domain, or several. All model TAAs contain a YLH, except for EibA where the YLH is not present. YpYadA has an "uptake region" at the beginning of the YLH. The Eibs group have an N-terminal region of unknown structures positioned before the YLH. The percentage of sequence similarity in the passenger between different TAA pairs is indicated.

3.2 Optimization of sedimentation assay

Before the sedimentation assay carried out for measuring the (auto- and co-) aggregation, the conditions for the experiments had to be optimized.

3.2.1 Optimization of IPTG concentration

A series of IPTG concentrations were tested for having a good protein expression (0.1, 0.5 and 1 mM) in both LB and the defined, clear medium PA-0.5G (Studier, 2005).
0.2% glucose has been included to the both media in order to repress background expression of TAAs and fluorescent proteins. An un-induced sample was also included as the negative control. So, the IPTG within the concentrations mentioned above had been added to the bacterial cells, and once they started the co-expression of the proteins, we observed the fluorescent protein which is then measured using the plate reader.
For performing the experiment, we tested the induction of the production (TAAs + fluorescent protein (sfGFP and mCherry). For that two-representative samples were chosen for test. One was EibA/sfGFP and the other was EibD/mCherry

Based on the results, (**Figure 10)** the fluorescence of the cultures induced by IPTG increased gradually with time, while the fluorescence of the un-induced samples remained at background levels. The fluorescent protein expression increased similarly with time at both IPTG concentrations (0.5 and 1 mM) and in both media. The fluorescence in these conditions was higher than with the 0.1 mM IPTG concentration. The expression of fluorescent proteins was much better in LB medium than in PA-0.5G medium. Although there was no difference in fluorescence between the IPTG concentrations of 0.5 and 1 mM, the 0.5 mM IPTG was chosen for expression in the aggregation assays because we did not want excessive over-expression of TAAs. As outer membrane proteins, their over-production might harm the bacterial cells (Dvorak *et al.*, 2015).

After inducing, the bacterial culture was left to grow for 2.5 h; the reason behind this was that mCherry expression and development started after almost 1.5hr of induction but for full chromophore development more time was needed (Ai *et al.*, 2014), in contrast to sfGFP (Roberts *et al.*, 2016) (**Figure 10 B)**.

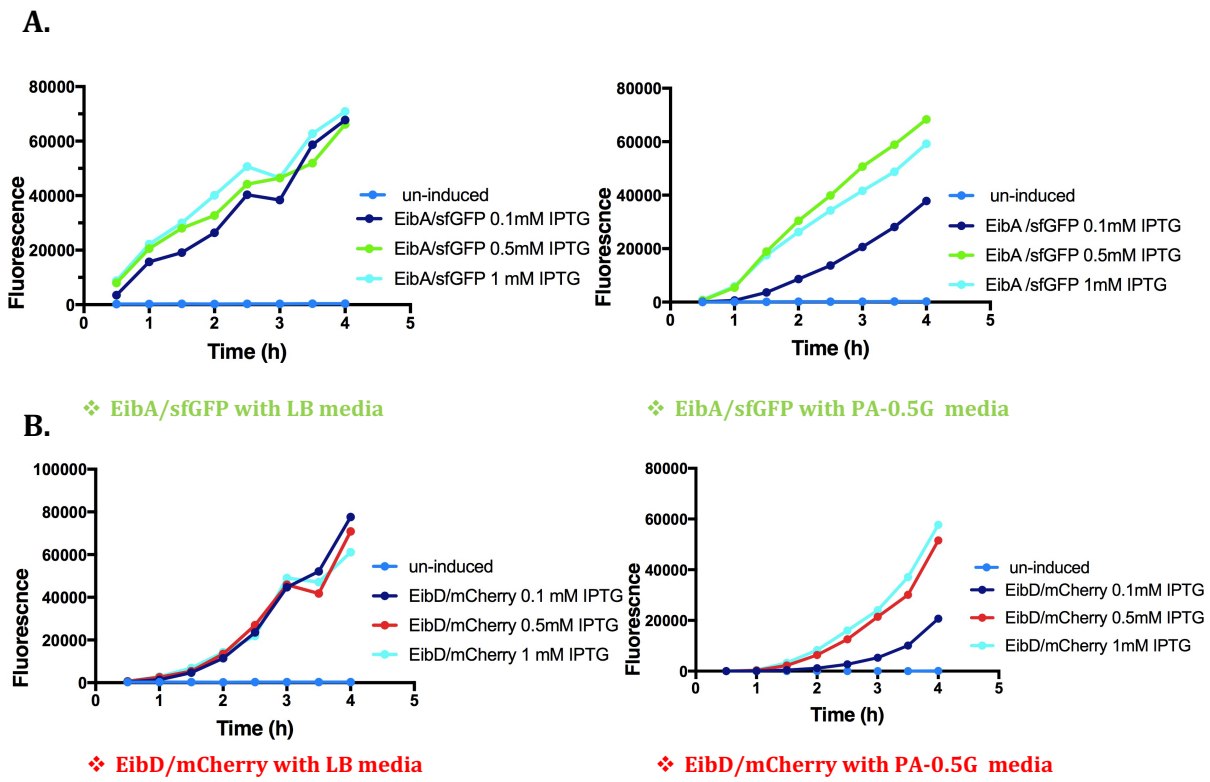


Figure 10. optimization of IPTG concentrations for aggregation assays in LB and PA-0.5G clear media. Two representative constructs were chosen for optimizing IPTG concentration, EibA/sfGFP (**A**) and EibD/mCherry with one biological replicate (**B**). For both, the fluorescence of cultures was measured using a plate reader at the excitation and emission wavelengths 483 and 510 nm, respectively for sfGFP, while for mCherry the excitation and emission wavelengths used were 580 and 610 nm, respectively.

3.2.2 Optimization of media

Another optimization step was the choice of growth medium. We tested two different media, LB and PA-0.5G defined medium (**Figure 11 A**). The reason we tested PA-0.5G, which is a clear medium, was to test a simple way for measuring the fluorescence by plate reader. Thus, rather than washing the cells as with LB that gives the background in the measurement, I could directly measure the fluorescence from the culture, which would save time and effort.

The LB media gives background in the fluorescence reading, and I had to wash the cells with PBS before measuring the fluorescence due to high background fluorescence caused by LB(Milbredt & Waldminghaus, 2017; Waters, 2009).

For this, two representative constructs were chosen, one representative with sfGFP (EibA/sfGFP) and the other with mCherry (EibD/mCherry). Based on results shown in (**Figure 11 B**) we chose the LB medium for our experiments because the expression of the fluorescent protein was higher than in PA-0.5G. In addition, the latter would still require washing the samples before reading the fluorescence; therefore, using the PA-0.5G medium would not have had any advantage.

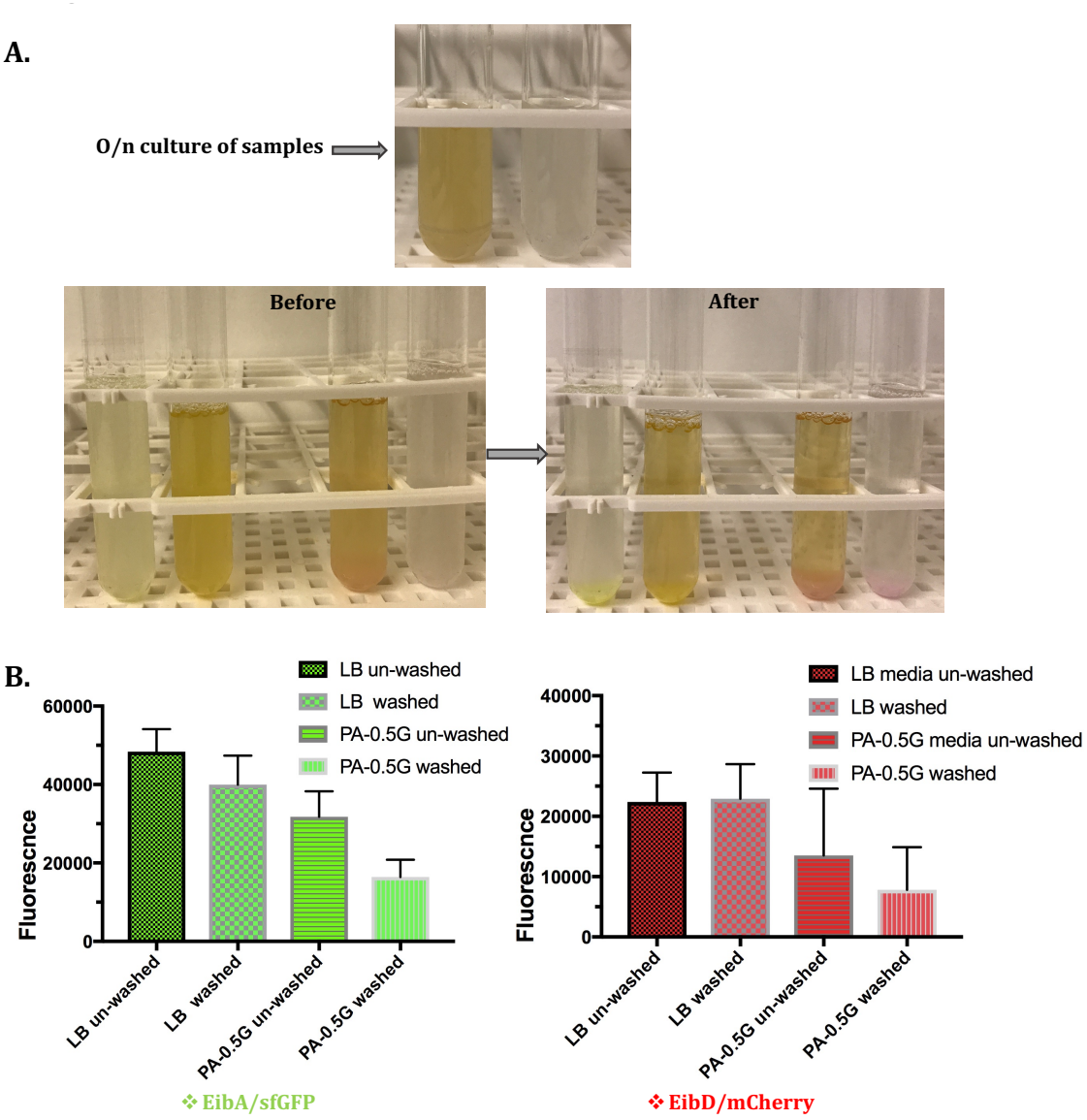


Figure 11. Optimization of media for autoaggregation assays.
(**A**) Bacterial cultures co-expressing a TAA and fluorescent protein before and after induction, left in PA-0.5G medium and right in LB medium. The production of fluorescent proteins is evident in the cultures. Two representative constructs, EibA/sfGFP and EibD/mCherry, were used. (**B**) Bar graphs representing fluorescence for both media. Left panel: EibA samples in both media showing that there is a reduction in fluorescence for washed samples in comparison with un-washed samples. The fluorescence was measured using a plate reader at the excitation and emission wavelengths 483/510 nm for sfGFP, and 580/610 nm for mCherry. The bars display the mean and SD of three biological replicates.

3.2.3 Bacterial sedimentation assay for un-induced samples

Another control experiment was done in order to check whether the constructs autoaggregate before induction with 0.5 mM IPTG. To this end, all the bacterial cultures were grown o/n and diluted the day after, then grown until the OD_{600} value reached 0.5. After that, the bacterial cultures were incubated for 30 min at 37°C without adding IPTG, then the OD_{600} was measured (**Figure 6** in materials and methods).
The results shown in (**Figure 12**) demonstrate that none of the un-induced bacterial cultures were clumping in the absence of inducer. In some of the samples that I took images of I noticed that there were some small clumps of bacteria, but similar small clumps were also present in the controls samples.
As the OD_{600} value of the cultures showed no dramatic reduction during the course of the experiment (3.5 h), I concluded that the small clumps would not significantly affect the outcome of later experiments.

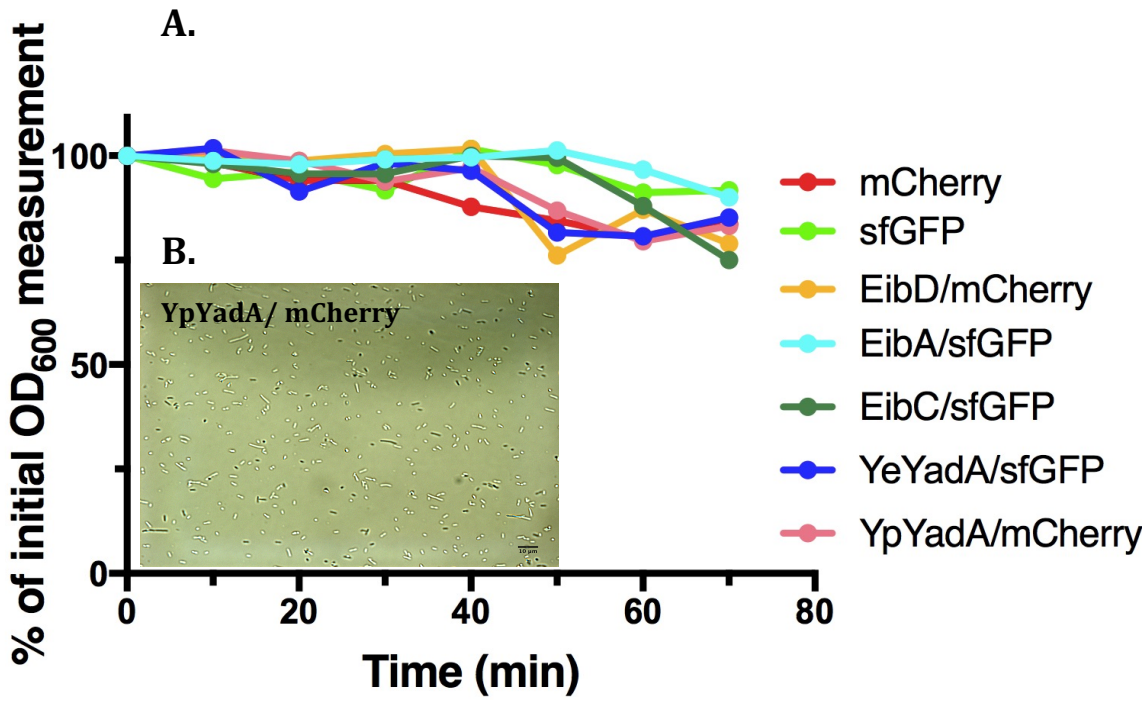


Figure 12. Bacterial cultures containing expression constructs do not aggregate without induction. The constructs made in this study were transformed in to BL21 (DE3) Gold. (**A**) None of the un-induced samples aggregated in a sedimentation assay with one biological replicate (**B**) Phase contrast micrograph for YpYadA/mCherry showing no formation of clumps. Scale bar=10μm.

3.3 Bacterial aggregation

3.3.1 Bacterial autoaggregation mediated by TAAs

To examine whether all TAAs mediate autoaggregation as a control experiment, sedimentation assays were performed using the all the optimized conditions (**Section 3.2**). For each individual sample that have been made in this study (**Section 2.1.3**) and with three biological replicates, the experiment has been performed by growing bacterial cultures o/n and then diluting the day after and growing until OD_{600} reached 0.5 (**Section 2.3.1** in materials and methods).
The fluorescence at the top of the cultures was measured at regular intervals and the data were plotted (**Figure 13 A**).

The results show that the control samples which are the bacteria without TAAs remain in suspension, while the constructs carrying a TAA autoaggregated (**Figure 13 B**). In two representative samples shown in (**Figure 13**), the percentage of initial fluorescence decreased quickly, and also the aggregation of bacterial cells was seen by microscopy, and it was also observed for all other constructs with TAAs (data not shown). The TAAs in general produce more or less spherical aggregates (**Figure 13 B**), but YpYadA tends to form elongated, sausage-like aggregates (**Figure 13 D**) and (**Figure 14 B**). This demonstrates that all TAAs used in this study mediated autoaggregation, whereas cells expressing only fluorescent proteins did not (**Figure 13 C**).

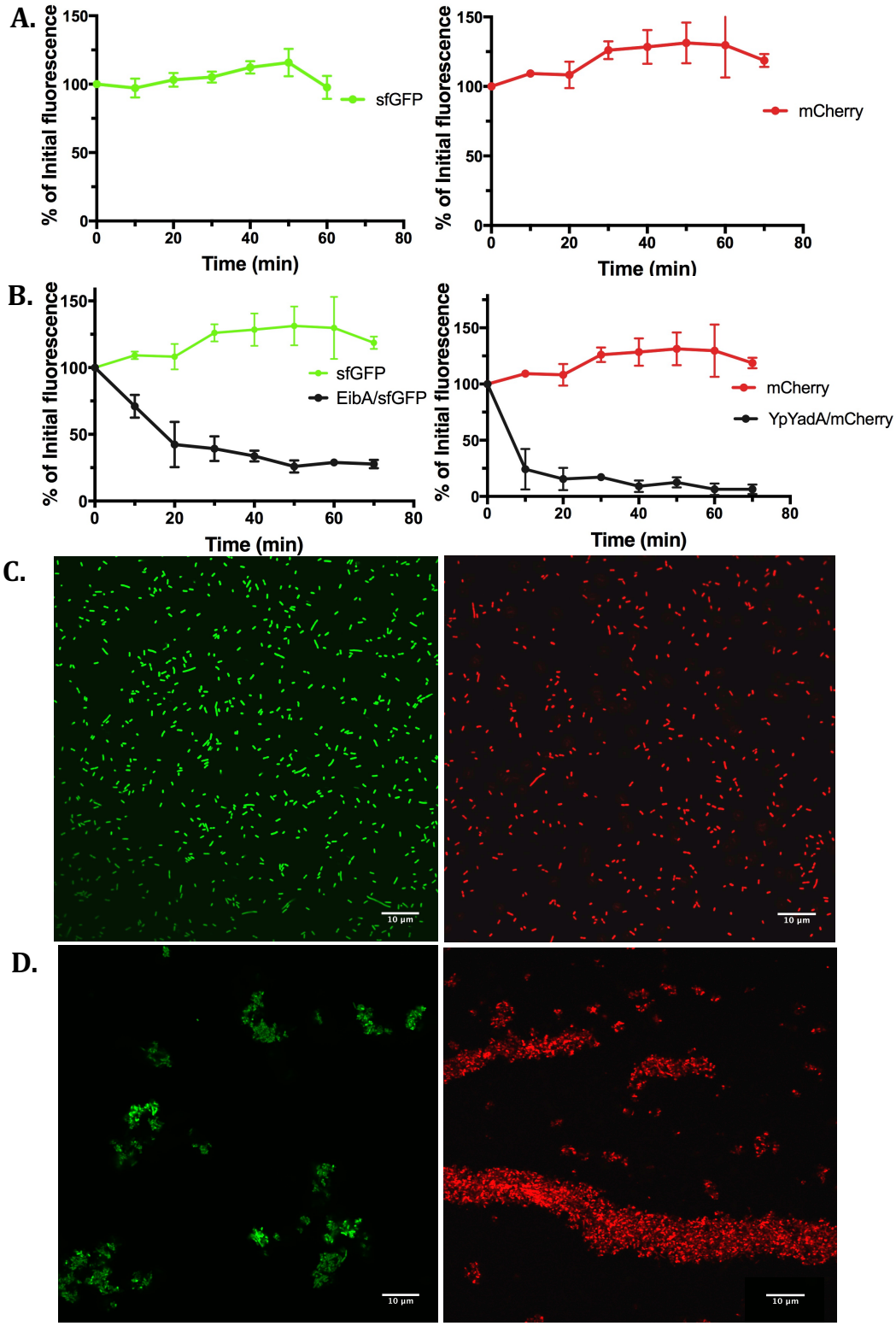


Figure 13. Autoaggregation mediated by TAAs as measured by sedimentation assays. (A) The graphs to the left represent the reduction of initial fluorescence for representative samples. The two first, sfGFP and mCherry, are control samples without TAAs. The results show that the cultures are not autoaggregating and still remain in suspension, while the two other constructs (**B**) EibA/sfGFP and YpYadA/mCherry, autoaggregate as shown by the rapid decrease of fluorescence at the top of the cultures. The graph displays the mean and SD of three biological replicates **(C)** CSLM of the autoaggregated samples. The two first micrographs show control samples sfGFP in green and mCherry in red, which are present as individual cells. (**D**) Autoaggregation mediated by TAAs is clear in the two displayed micrographs showing EibA/sfGFP formed clumps whereas YpYadA/mCherry formed elongated, a sausage-like aggregates. Scale bar=10µm.

3.3.2 TAAs mediate co-aggregation based on sequence similarity and Image analysis of TAAs

To measure co-aggregation, I investigated the interaction between cells expressing different TAAs and fluorescent proteins by sedimentation assays and microscopy. The combinations of TAAs and fluorescent proteins I used are given in (**Table 9**)**.** For preforming the experiment mixed bacterial culture were grown o/n and diluted the day after, then grown until the OD_{600} value reached 0.5. An equal amount of each bacteria from two cultures were mixed together and then induced by using 0.5 mM IPTG. After induction, the fluorescence measured at regular intervals using a plate reader and the data were plotted (**Figure 14**)**.**

Table 9. Lists of mixed groups of TAAs with controls group. The green colour represents the construct with sfGFP and the red colour represent the construct with mCherry. The control groups are either with TAAs alone or two mixed TAA of the same type or a plasmid lack TAA.

Mixed control groups	Groups of mixed TAAs
EibA + mCherry	YeYadA + YpYadA
EibC + mCherry	EibA + YpYadA
EibD + EibD	
sfGFP + EibD	YeYadA + EibD
sfGFP + mCherry	EibA + EibD
sfGFP + YpYadA	EibC + EibD
YeYadA + mCherry	EibC + YpYadA
YeYadA + YeYadA	

After preforming the sedimentation assays for co-aggregation, at the end of experiment I took some of the sediment from the samples for CLSM (**Figure 14 B**). The images were processed first by using imageJ software and implementing a script to calculate the association index. The association index is given as the percentage of green bacteria (expressing sfGFP) that have a red bacterium (producing mCherry) within 15 pixels, and vice versa. The results are given as the mean of six random fields from two biologicals replicates (**Table 10**).

Based on the association index calculated, I classified the samples into three categories of co-aggregation: -

1. Completely mixed as in the case of EibD/mCherry + EibD/sfGFP at range of 30% and above.

2. Intermediately mixed, like EibD/mCherry + EibC/sfGFP and EibD/mCherry + YeYadA/sfGFP at range of 10% and above.

3. Non-mixed (exclusion) which have a lower association index than background. Examples are EibA/sfGFP + EibD/mCherry and YeYadA/sfGFP YpYadA/mCherry. YpYadA/mCherry + EibA/sfGFP at range below 10%.

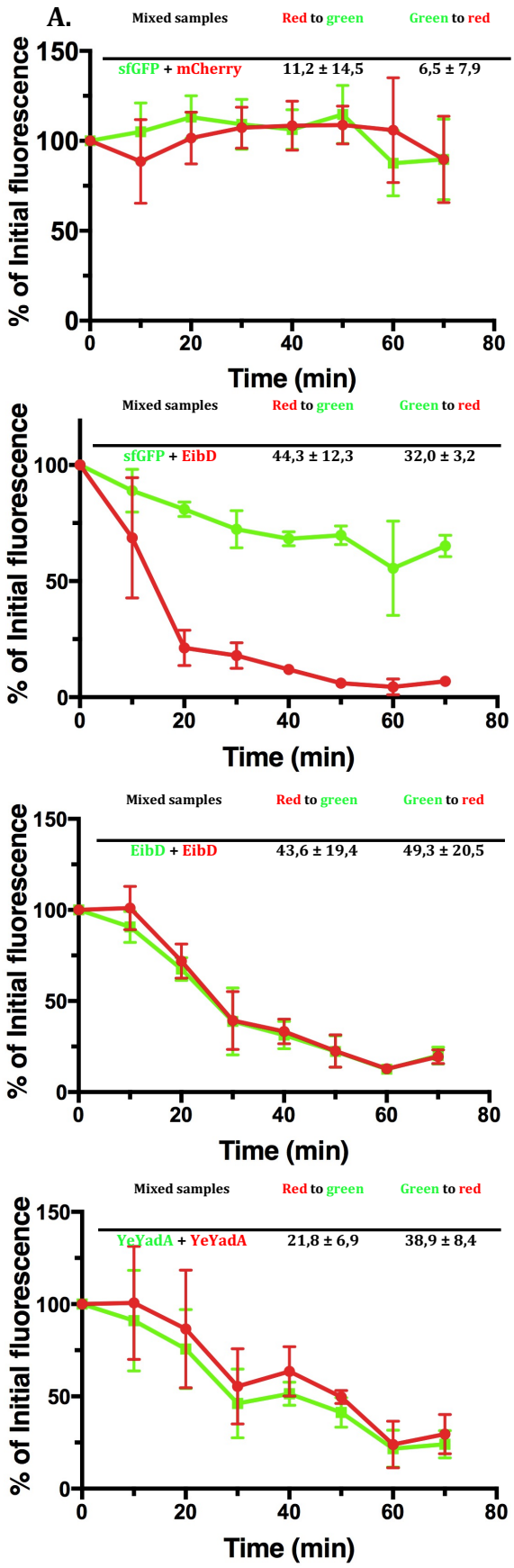
The first one which was the completely mixed (homotypic interaction as in the case of EibD/sfGFP + EibD/mCherry a control sample, and YeYadA/sfGFP + YeYadA/mCherry. The second category is the intermediate mixing as in the case of EibC + EibD, which have the high overall association index 29.3%. This can be explained by the high sequence similarity between these proteins (89%).

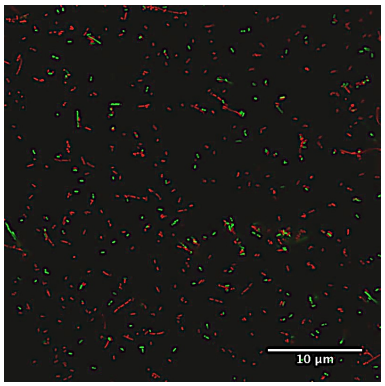
Another example for the second category or the intermediate mixing is EibD + YeYadA. This shows for the first time that two different TAAs can also co-aggregate together and the overall association index was 17.8%, although the sequence similarity in the passenger was only 25%, but they showed a heterotypic interaction mediated by two different TAAs.

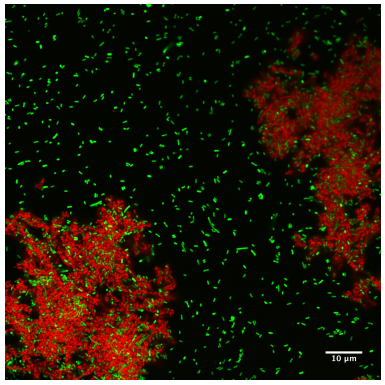
Although the interaction, exemplified by EibD/mCherry + sfGFP (control sample), showed a high overall association index (38.6%), we did not include it in the first or the second category of interaction. We expected the sample to have a low association index, but this was not the case. The reasons for the high association index here is most likely due to the way the index is calculated because of the limitation of the (current) script. The program cannot distinguish the difference sizes of bacterial aggregates, and even in the micrograph (**Figure 14 B**) we could see that the sfGFP sample is distributed evenly, with some cells embedded in the TAA-expressing aggregates. The same issue was seen for the samples that with the mCherry control (micrographs not shown).

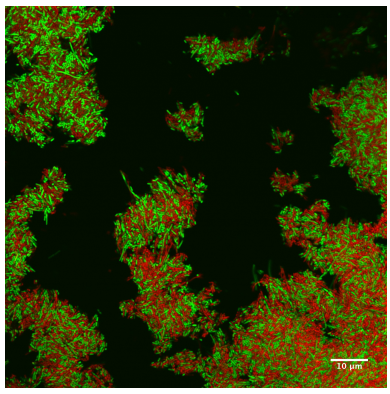
We can also see another deviation from expected in the case of YpYadA + EibA. In the micrograph shown in **Figure 14B**, there appears to be an intermediate interaction between EibA and YpYadA, but the association index for them predicts exclusion. Again, we suggest this is due to limitation of the script used for quantification.

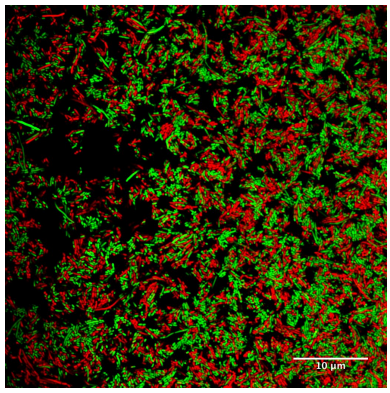
For the Third category, although the sequence similarity was as high as 69% in the case of YeYadA + YpYadA, and we expected that these two TAAs will co-aggregate, but both mixed TAAs did not co-aggregate, i.e. they excluded each other. The overall association index was 6.9% in case YeYadA + YpYadA. Similarly, the sequence similarity for EibA + EibD pair was 37%, but the overall association index was 8.1%, which places this pair also in the non-mixed category. The main difference between EibA and EibD is the lack of the YLH in EibA, whereas the big difference between YeYadA and YpYadA is the uptake region found only in YpYadA. The other reason is that the uptake region might cover the YLH and prevent binding of both yadA, because the uptake region was also responsible for changing the yadA main function for binding to fibronectin rather than collagen (Heise & Dersch, 2006)

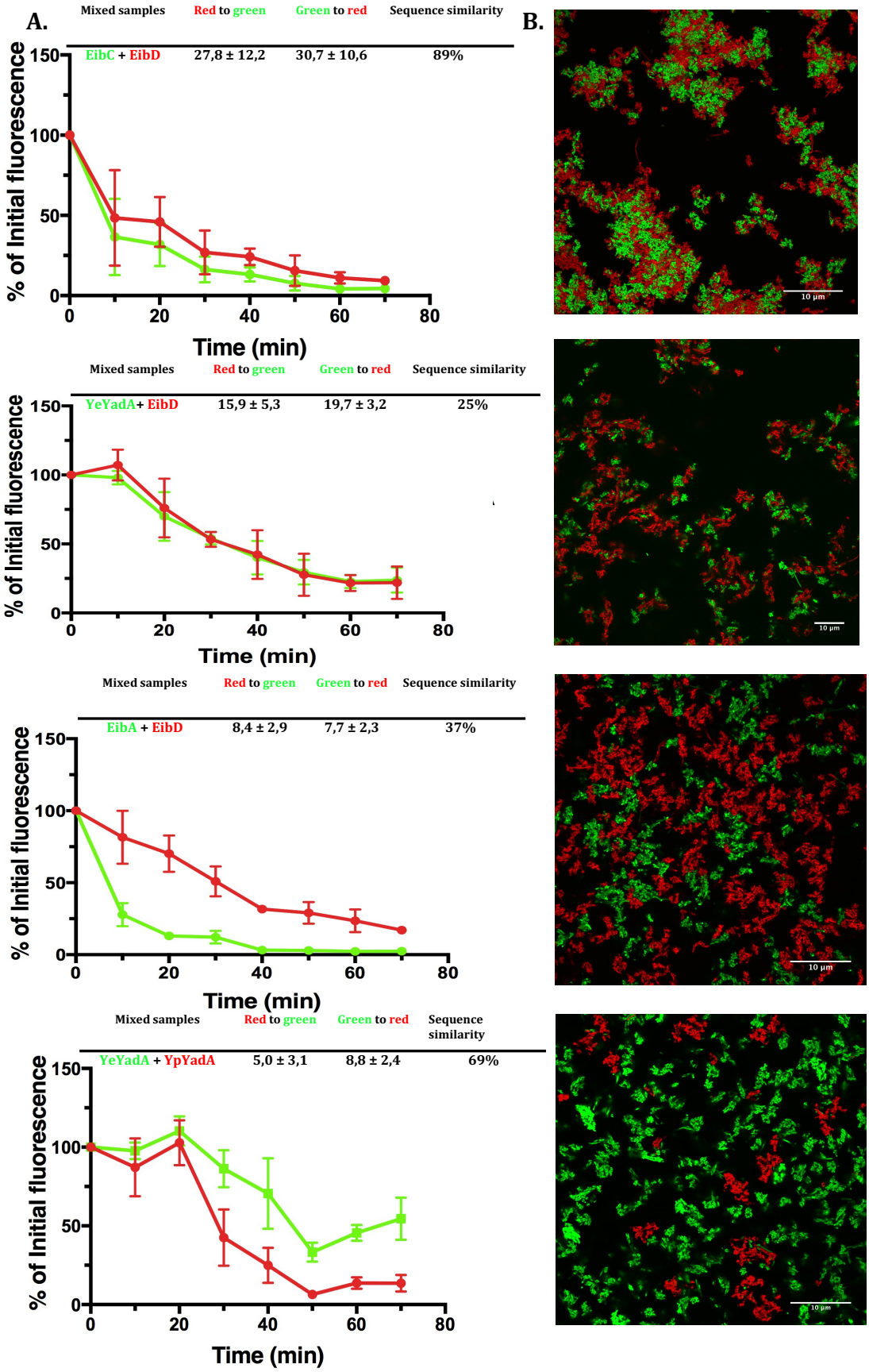

A.
Mixed samples
Red to green
Green to red
sfGFP + mCherry
11,2 ± 14,5
6,5 ± 7,9
% of Initial fluorescence
Time (min)
sfGFP + EibD
44,3 ± 12,3
32,0 ± 3,2
EibD + EibD
43,6 ± 19,4
49,3 ± 20,5
YeYadA + YeYadA
21,8 ± 6,9
38,9 ± 8,4


B.
10 µm


10 µm


10 µm


10 µm


A.
B.
Mixed samples
Red to green
Green to red
Sequence similarity
EibC + EibD
27,8 ± 12,2
30,7 ± 10,6
89%
YeYadA+ EibD
15,9 ± 5,3
19,7 ± 3,2
25%
EibA + EibD
8,4 ± 2,9
7,7 ± 2,3
37%
YeYadA + YpYadA
5,0 ± 3,1
8,8 ± 2,4
69%
% of Initial fluorescence
Time (min)
0
20
40
60
80
50
100
150
10 µm

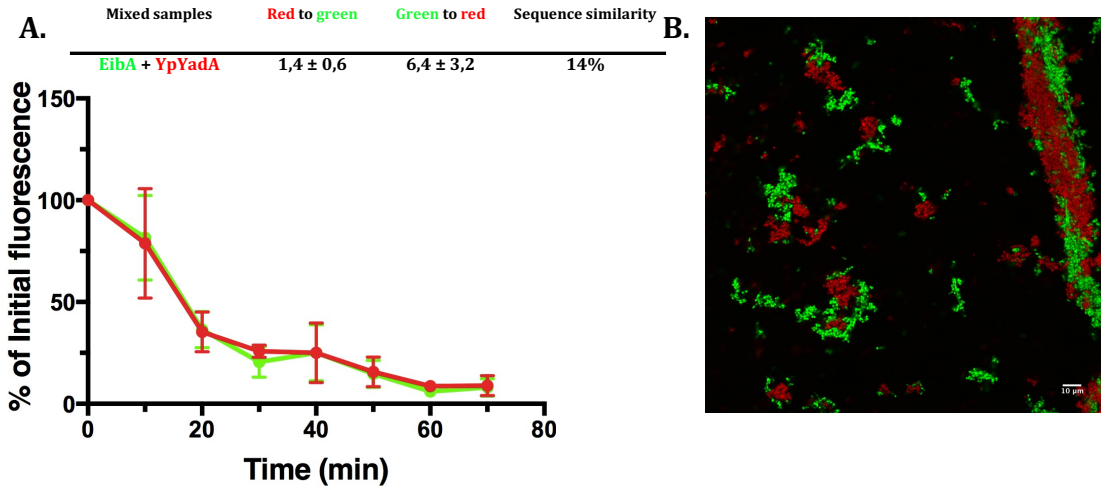


Figure 14. Co-aggregation using sedimentation assays. Representative results are shown: (**A**) TAA expression causes bacteria to aggregate and sediment, whereas control cells with only the fluorescent protein remain in suspension. The graph displays the mean and SD of three biological replicates (**B**) CSLM shows that the bacteria are aggregating and clumping together in case of EibD/mCherry with YeYadA/sfGFP, while some other exclude each other such as EibD/mCherry with EibA/sfGFP. Scale bar=10µm

Table 10. Association of bacteria expressing different TAAs. The values are the association index calculated using a script implemented in imageJ. This is given as the mean ± SD from six random fields taken using CSLM for each sample representing two biological replicates.
Note: The overall association index is the average of the two association indices (red to green and green to red) calculated from the data for each sample.

Mixed Sample	Association index				
	Red to green	Green to red	Overall	Category	Similarity
EibA + EibD	8,4 ± 2,9	7,7 ± 2,3	8,1	Exclusion	37%
EibA + YpYadA	1,4 ± 0,6	6,4 ± 3,2	3,9	Exclusion	14%
EibA + mCherry	2,4 ± 1,2	28,5 ± 22,5	15,6	Background	-
EibC + EibD	27,8 ± 12,2	30,7 ± 10,6	29,3	Intermediate	89%
EibC + YpYadA	12,1 ± 2,7	15,7 ± 18,5	13,9	Intermediate	-
EibC + mCherry	8,5 ± 4,2	20,6 ± 11,8	14,6	Background	-
EibD + EibD	43,6 ± 19,4	49,3 ± 20,5	46,5	mixed	-
sfGFP + EibD	44,3 ± 12,3	32,0 ± 3,2	38,6	Background	-
sfGFP + mCherry	11,2 ± 14,5	6,5 ± 7,9	8,9	Exclusion	-
sfGFP + YpYadA	13,0 ± 5,3	29,4 ± 13,2	21,2	Background	-
YeYadA + YeYadA	21,8 ± 6,9	38,9 ± 8,4	30,4	mixed	-
YeYadA + EibD	15,9 ± 5,3	19,7 ± 3,2	17,8	Intermediate	25%
YeYadA + YpYadA	5,0 ± 3,1	8,8 ± 2,4	6,9	Exclusion	69%
YeYadA + mCherry	5,0 ± 1,7	17,2 ± 7,9	11,1	Background	-

3.4 Biofilm formation

3.4.1 TAAs mediate the formation of biofilm on different Surfaces

To investigate whether TAAs used in this study will mediate the formation of biofilm on different surfaces. Each individual construct (without mixing)
were grown o/n and diluted the day after, then grown until the OD_{600} value reached 0.5. Then, biofilm cultures were prepared by adding 3 µl of bacterial culture (un mixed) with 3 ml of autoinduction media in to two different surfaces (glass and Polystyrene) (**Section 2.6** in materials and methods). Autoinduction medium (**Appendix 4** for media) was used for biofilm formation because it gives higher cell density compared to IPTG as the inducing agent and for its simplicity of use, as the inducer does not need to be added separately.

The results (**Figure 15**) show that EibA formed biofilm on the glass surface more than the other TAAs. For testing biofilm formation on polystyrene, I set up cultures both with and without agitation in order to know whether there is a difference between these two conditions, and the YpYadA formed more biofilm on polystyrene (**Figure 15 B and C**) than the others. EibD formed less biofilm and. In general, the results showed that all cultures (including the control) formed more biofilm under static conditions, but the differences (particularly relating to the control) were more pronounced with gentle agitation.

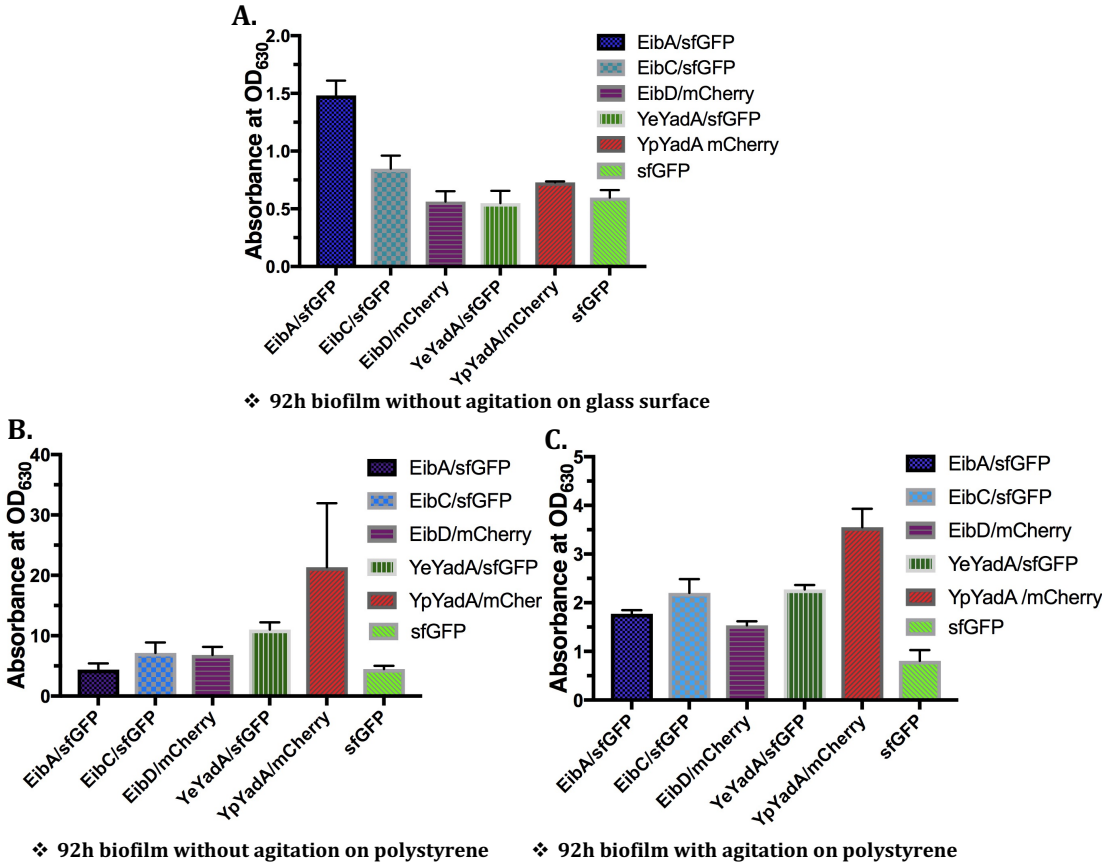


Figure 15. Biofilm formation mediated by TAAs on different surfaces. Biofilms were formed for 92 h on different surfaces and then stained with crystal violet within three biological replicates. For quantification, absorbance's were recorded at 630 nm. (**A**) Biofilm formation on glass surface under static conditions, which shows that EibA/sfGFP forms the most biofilm. While, on polystyrene, (**B**) showing that the YpYadA/mCherry forming the most biofilm on static condition. While, the same sample on the (**C**) left shows that it forms less biofilm within agitation at 40 rpm/min.

3.4.2 Biofilm formation by mixed populations expressing different TAAs

Based on the results from co-aggregation assays, some of combinations of interest were chosen (**Table 11**) for forming mixed biofilm (see **Section 2.7** in materials and methods). After 92 h of incubation, 3D images (**Figure 16**) were taken for all samples using an Andor Dragonfly confocal microscope, which was used because it took images very fast: each sample needed only few seconds. It provides high resolution, and also I did not destroy or damage the structure of biofilm by discarding the media from the plates with biofilm and washing them with 1xPBS, because when I did that I noticed that the structure of the biofilm was affected.
The 3D maximum projection was processed in Imaris Suite (Bitplane). For the 3D object-based colocalization and Imaris XTension spot co-localization, used to analyse the interaction with the neighbouring bacteria. Within a distance of 1.15 µm (**Section2.5.2** in materials and methods).

Based on the association index calculated, I classified the samples into two categories of biofilms: -

1. Completely mixed biofilm as in the case of EibD/mCherry + EibD/sfGFP, YeYadA/ mCherry + YeYadA/sfGFP (see **Table12**) at range of 15 % and above.

2. Non-mixed (segregated) biofilm, like which have a lower association index than background. Examples are EibA/sfGFP + EibD/mCherry and YpYadA/mCherry + EibA/sfGFP at range below 10%.

Based on the results, (**Table 12**) the association index of YeYadA + EibD was higher than the index for other mixed cultures, which indicates that there is still co-aggregation between the two strains in the biofilm and forming the mixed biofilm. What was interesting here that the two-mixed sample from YadA group YeYadA + YpYadA are also showed a kind of interaction and formed a mixed biofilm, while the two-mixed sample form Eibs group the EibD + EibA were still excluding each other and formed a segregate domain.

Table11. List of mixed bacterial cultures expressing TAAs and controls used for biofilm assay

Mixed control group	Group of mixed TAAs
EibD + EibD	EibA + EibD
sfGFP + EibD	YeYadA + EibD
YeYadA + YeYadA	YeYadA + YpYadA
	EibA + YpYadA

Table 12. The association index of differently coloured bacteria in biofilm formation assays. Samples were analysed using Imaris XTension spot co-localization within distance of 1.15 µm. This is given as the mean ± SD from random fields taken using an Andor Dragonfly microscope for each sample representing three biological replicates.

Mixed Sample	Association index			
	Red to green	Green to red	Overall	Category
EibA + EibD	4,2 ± 1,2	3,8 ± 1,5	4	Non-mixed
EibA + YpYadA	7,3 ± 3,6	9,8 ± 2,3	8,6	Non-mixed
EibD + EibD	44,2 ± 2,2	32,3 ± 2,0	38,3	Mixed
sfGFP + EibD	21,1 ± 18,2	31,5 ± 21,5	26,3	Background
YeYadA + YeYadA	18,9 ± 0,9	19,0 ± 3,8	19,0	Mixed
YeYadA + EibD	29,7 ± 1,7	12,8 ± 1,2	21,3	Mixed
YeYadA + YpYadA	18,3 ± 9,7	18,4 ± 4,4	18,4	Mixed

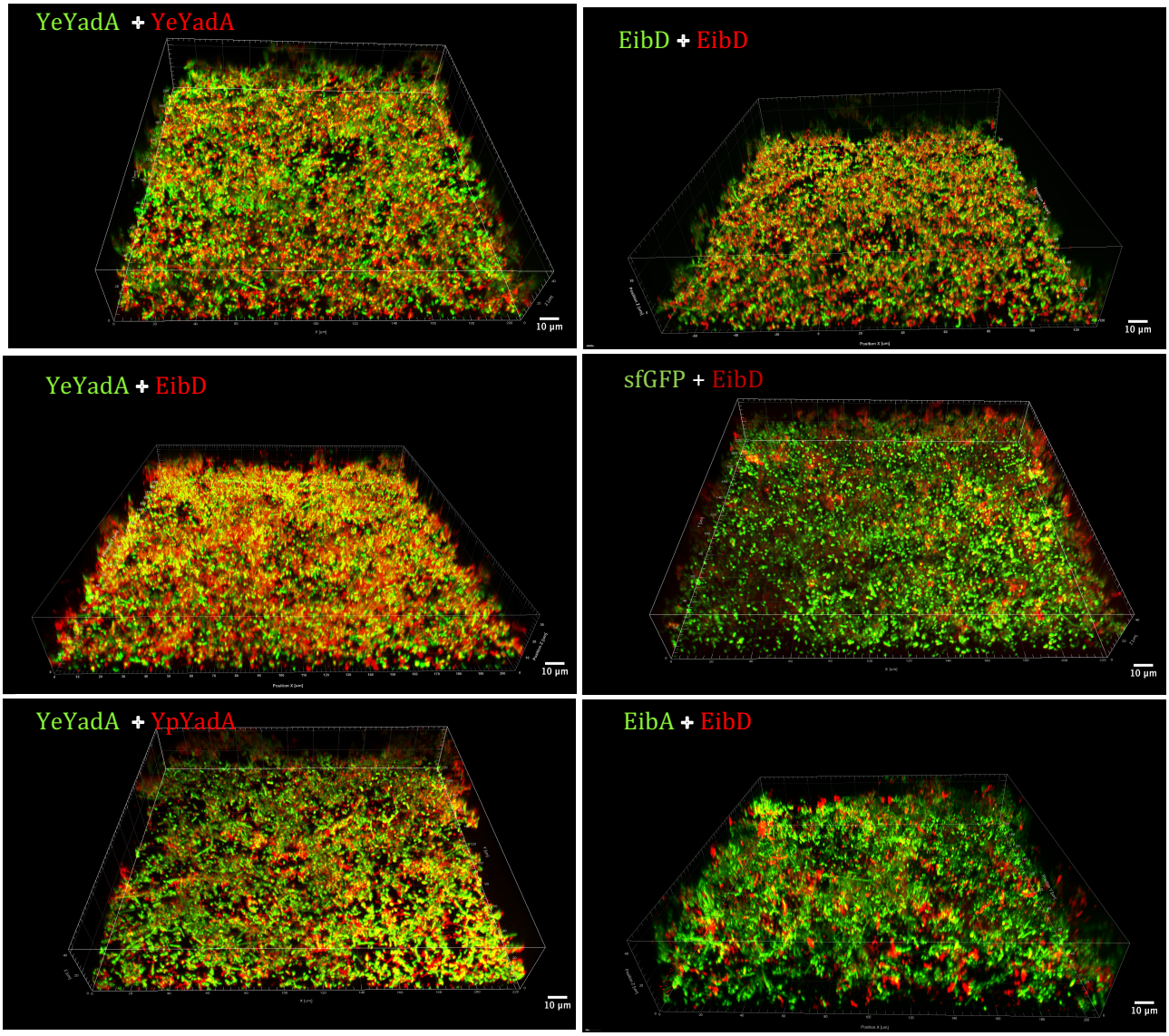


Figure 16. 92hr biofilm formed from static mixed bacteria expressing TAAs. 3D maximum projection was processed in Imaris Suite. The images were taken by using Andor dragonfly microscope shows that the bacteria are mixed together in the biofilm matrix. Scale bar=10µm.

3.5 Mutagenesis

In order to know which residues or domains in TAAs are responsible for mediating autoaggregation, one step site-directed plasmid mutagenesis was performed in order to abrogate the autoaggregation properties for TAAs. A series of point mutations was designed based on the crystal structure (PDB ID:1P9H) of YeYadA head domain (**Figure 18**). Our hypothesis was that residues make crystal contact will also be involved in autoaggregation. Some of these interactions were polar interactions, for example Q124 of one monomer interacts with Q124 of another monomer and we assumed changing Q124 into A(alanine) will disrupt the autoaggregation, but the results showed that either single point mutations in the head domain of YeYadA did not affect the sedimentation of the bacteria; the bacterial cells expressing YadA variants with a single point mutation still aggregated (**Figure 18 A**). However, there was a very slight difference between the different mutants. For example, I94A, E182A, Q124A and L110A showed a delay in aggregation and settling of the bacteria. For that reason, these residues were combined in to single construct (**Figure 18 B**).

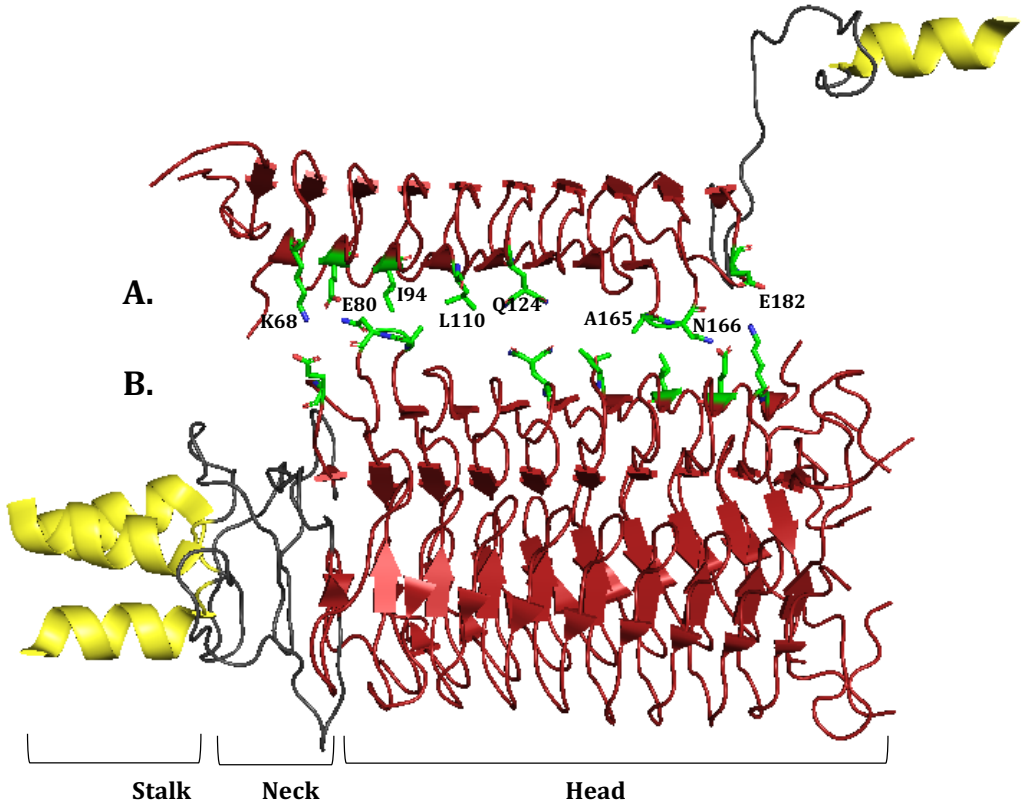


Figure 17. Mutagenesis of surface residues of YadA. The model structure of YeYadA head domain the figure was made using PyMol (PDB ID:1P9H) showing (**A**) residues that make crystal contact with (**B**) another YadA monomer.

In order to abolish autoaggregation. A double and triple mutations in the head domain of YeYadA were made by combining some of the single mutations in one plasmid. However, no major effects were seen for these (**Figure 18 B**).
In addition to point mutations, I also deleted the whole head domain of YeYadA. To test which domains, affect autoaggregation in EibD, we deleted the N-terminal Eib region, and

separately the YLH domain was also deleted. For YpYadA, I deleted the uptake region (**Figure 9**) to test whether this was important for autoaggregation.
ALL these constructs mentioned above (deletion of domain) were tested for autoaggregation using the sedimentation assay. However, we did not observe any changes in the autoaggregation behaviour: the bacterial cells were still aggregating and sedimenting (**Figure 18 C**).

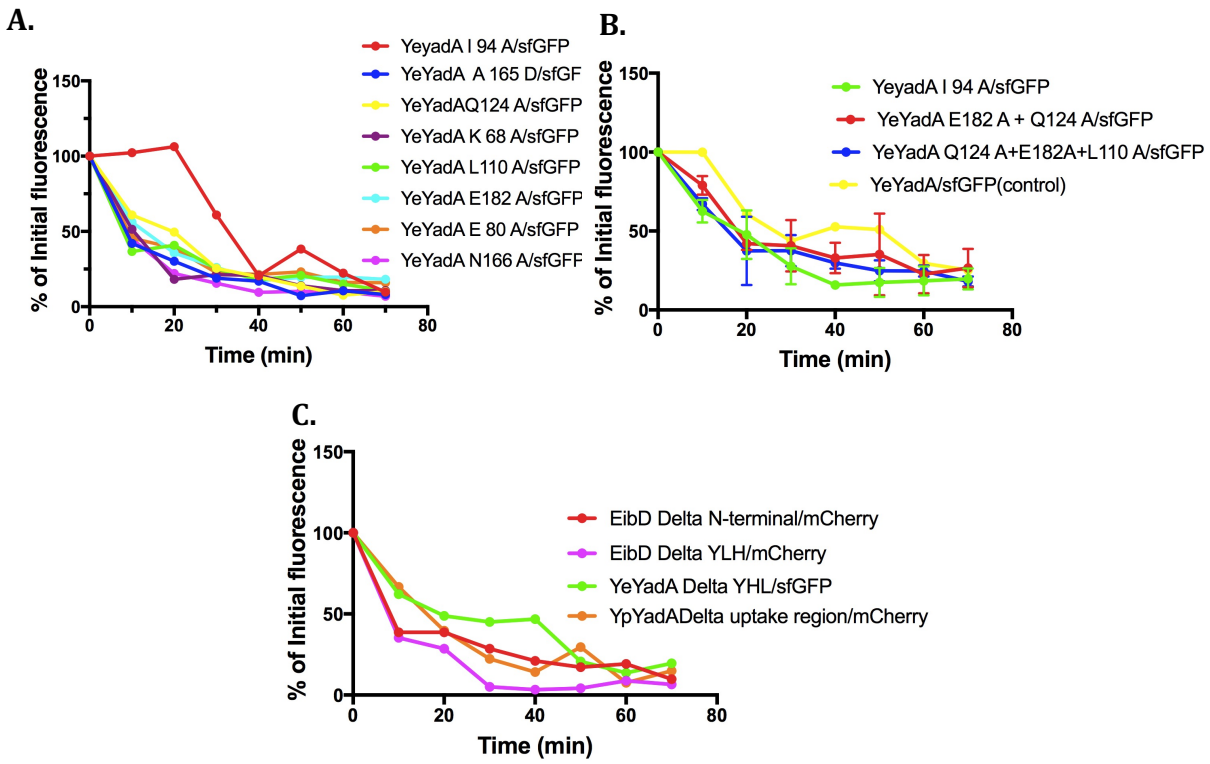


Figure (18) Sedimentation assay for mutant constructs (**A**)All single point mutation in the YLH domain of YeYadA/sfGFP were tested for autoaggregation with one biological replicate, (**B**) Based on results from (**A**) double and triple mutations were also made in YeYadA/sfGFP YLH domain and tested for autoaggregation. The single point mutation I94A and the wild-type were controls. The bars represent the mean and SD for three biological replicates. (**C**) Autoaggregation of constructs with deletions of the whole YLH domains in YeYadA and in EibD was tested. Also, the N-terminal domain of EibD and the uptake region of YpYadA were deleted and the autoaggregation properties tested with one biological replicate.

4 Discussion

The aim of this study was to characterize the co-aggregation properties of trimeric autotransporter adhesins. We used YeYadA, from *Yersinia enterocolitica*, which is a protein that has been the subject of many studies, and YpYadA from *Y. pseudotuberculosis*. In addition, three different proteins from the Eib group, EibA, C, and D, from *E. coli* were used.
Both groups of TAAs are known to promote autoaggregation, seen easily by the formation of clumps of bacteria that settle down at the bottom of culture tubes. The interaction leading to autoaggregation is homotypic, but it was not known if different TAAs mediate heterotypic interactions when mixed together. Therefore, we addressed this question by co-expressing a fluorescent marker and a TAA in the same cell. We showed that all the TAAs tested in this study mediate autoaggregation. This includes EibA and EibC, for which autoaggregation had not been demonstrated before. Furthermore, no aggregation was seen in the absence of inducer. I optimized conditions and proceeded to test co-aggregation of TAA-expression bacteria.

4.1 TAAs mediate co-aggregation

Before performing co-aggregation experiments, several optimization steps had to be taken in order to get the best expression of protein and measurement conditions. Importantly, we also checked that the constructs made were working properly. As a control experiment, we wanted to know if the TAAs are aggregating before adding the IPTG. Our experiment showed that without induction none of the strains aggregated.

To check that all constructs mediated autoaggregation, sedimentation assays were performed. In all the samples the fluorescence at the top of the cultures decreased over time (representative samples shown in **Figure 13**), which means that the TAA proteins were expressed and that all of them mediated autoaggregation, as expected (Heise & Dersch, 2006; Hoiczyk *et al.*, 2000; Leo *et al.*, 2011; Skurnik *et al.*, 1984). In contrast, the cells in the control cultures lacking TAAs remained in suspension.

For performing co-aggregation assays, many studies (Cheng *et al.*, 2014; Rosen *et al.*, 2008) did the assay in a different way from our approach. In one way, aggregation is measured using a purely visual assay, where an equal volume of the bacterial cells was mixed and the mixed culture was vortexed and then let stand at room temperature. At specific time intervals, the tubes were mixed again for 10s and the investigators measured the co-aggregation by giving scores on a scale 0-4 for the degree of co-aggregation.
Another way to measure co-aggregation is a spectrophotometric assay, where the co-aggregating samples are prepared by the same method as above, but for measuring the turbidity of the supernatant was recorded at different intervals of time, after which the percentage of co-aggregation was calculated (Cheng *et al.*, 2014).

In our study, in order to determine the co-aggregation between different TAA-expressing bacteria, we mixed equal amounts of bacterial cells based on OD_{600} measurements, which is quantitative. The cultures were then induced after 30 min of incubation together. In all the mixed samples, TAA-containing bacteria aggregated, as shown by the reduction of fluorescence at the top of the cultures over time (**Figure 14 A**). So even though these TAA-

expressing bacteria were mixed together, they still aggregated. However, when a TAA-expressing culture was mixed with a control culture lacking a TAA (**Figure 14 A**), the fluorescence corresponding to the control samples remained close to 100%, which means that there was no cell-cell interaction or recognition happening between TAAs and control samples. Although we could see bacteria trapped in TAA-mediated aggregates in the micrographs (**Figure 14 B**), we do not think this is a specific interaction. Rather, control cells become embedded in the aggregates by chance.

The sedimentation assays of mixed cultures did not provide enough information to know whether the mixed TAA-expressing bacteria were co-aggregating or not. Therefore, we used CSLM to image the samples and the images were processed further using a script that calculates the association index of bacteria of a given colour to bacteria of the other colours.

Based on this analysis, we could see three different categories of co-aggregation. The categories have been determined based the range of the overall association index. The first category was the completely mixed as in the case of EibD/mCherry + EibD/sfGFP at range of 30% and above.
The second category was the intermediately mixed at range of 10% and above, in case of the EibD/mCherry + EibC/sfGFP, and they showed the higher overall association index than the others.
The degree of co-aggregation in this case is most likely correlated to the degree of the sequence similarity. Another example for the intermediate interaction is the EibD/mCherry + YeYadA/sfGFP. Although the sequence similarity is only 25% between EibD +YeYadA, but the results showed that there is a heterotypic interaction between them.
In general, the results analysis provided by the script was good, but it has also some limitations and gave some inconvenience results regarding to the control sample.
The high value of overall association index for sfGFP sample + EibD was not expected, because the sfGFP sample is distributed evenly, with some cells embedded in the TAA-expressing aggregates. So, the high value come from the fact that the script calculated each single bacterium of sfGFP sample around the aggregate of EibD.
Finally, the non-mixed (exclusion) which have a lower association index was EibA/sfGFP + EibD/mCherry, YeYadA/sfGFP + YpYadA/mCherry samples, and YpYadA/mCherry + EibA/sfGFP at range below 10%.

It was expected for YeYadA + YpYadA to co-aggregate together, because they have a high sequence similarity (69%), but they showed exclusion. The only difference between these TAAs is the uptake region, which is found in YpYadA but is not present in YeYadA. The EibA sample similarly excluded EibD, despite a relatively high similarity (37%). The main difference between EibA and EibD is the YLH region that is present in EibD, but not in the EibA.

4.2 Biofilm formation assay

Colonization and adherence are the key event in bacterial pathogenesis (Stones & Krachler, 2016). TAAs are known to mediate autoaggregation and, during this study, we also found that some of TAAs can also co-aggregate. Both the autoaggregation and the co-aggregation can lead the bacteria to form biofilm.
From the results of the co-aggregation assays, it was clear that some TAAs mediate co-aggregation, whereas others do not. Cell-cell interactions during biofilm production are crucial in determining the biofilm architecture (Colley *et al.*, 2016). Therefore, we were interested in finding out if the TAAs that co-aggregated together would mediate the formation of mixed biofilms or whether the different cell types would form microdomains within the biofilm.

First, I performed a quantitative assay for all the constructs to check biofilm formation. The resulting biofilms were stained with crystal violet. These experiments served a kind of biofilm optimization step and the results showed that all TAAs are able to form biofilm on both surfaces (**Section 3.4** in results). However, there were some differences, like EibA which formed more biofilm on glass than polystyrene. This suggests that EibA may make the cells more hydrophilic than the other TAAs, as glass is a hydrophilic surface. This might be explained by the absence of a YLH domain in this protein; for it has been suggested that the YLH domain confers hydrophobicity (Tamm *et al.*, 1993).

YpYadA mediated strongest biofilm formation among the samples that were incubated on polystyrene. This might be because YpYadA makes the bacteria more hydrophobic than the other TAAs. Consistent with our findings, it has also been reported that YpYadA mediates forming a cloudy biofilm-like layer on glass surfaces (Heise & Dersch, 2006).
Based in the results of analysis, the EibD + YeYadA sample that previously showed intermediate association in the co-aggregation experiments also formed mixed biofilm with an overall association index of 21.3%. From this, we can confirm that this the first time showing that the TAAs not only mediate co-aggregation but also mixed biofilm formation. However, TAAs might not be necessary for biofilms under these conditions because also the sfGFP control sample and EibD formed a similar mixed biofilm.

An interesting finding was that some of the TAAs like YpYadA with YeYadA showed co-aggregation forming a mixed biofilm. In spite of high sequence similarity (69%), YeYadA + YpYadA did not co-aggregate in the sedimentation assay. The micrographs show that they are excluding each other, as does the analysis using the script, where the pair is placed in the third category (exclusion). However, in the biofilm, the association index was relatively high (19%)**.** The reason for the difference is unclear; however, it might due to longer incubation and autoinduction conditions (Piwat *et al.*, 2015).
EibA + EibD were still excluding each other for the association index was 4% even in the biofilm. The reason for that is probably because EibA forms biofilm much better on glass than EibD, and as it does not aggregate with EibD it may simply prevent EibD-expressing cells from getting a foothold on the glass surface.
Although the association index of the EibD + sfGFP was high (26,3), but they could not form a mixed biofilm, because even BL21 alone (fluorescent protein (sfGFP)without TAAs) can form biofilm on glass (Hydrophilic surface); which means that sfGFP sample might exclude EibD to attach the glass surface, similarly as EibA + EibD did.

4.3 Mutagenesis

The final step in my project was to try to find residues and domains mediated autoaggregation in TAAs, which for me was the most interesting part because we want to abrogate this property.
Based on the literature (Hoiczyk *et al.*, 2000; A. Roggenkamp *et al.*, 2003), the head domain of YeYadA mediates the autoaggregation. However, for the EibD is not clear if the YLH domain mediates autoaggregation or not, and also EibA mediated the autoaggregation and it has not the YLH.
I made 8 single point mutations in the head domain for YeYadA and tested them using the sedimentation assay, but due to lack of time, the experiments were done with only one biological replicate. Nevertheless, the results showed all the mutants still mediated autoaggregation effect. However, four of these mutations (I94A, L110A, E182A, and Q124A) appeared to slow down the sedimentation somewhat.

I wanted to know whether the combination of these single mutation would abrogate the autoaggregation. Therefore, I made a combination of mutations by making a double (E 182 A + L110 A) and triple (Q124A + E182A + L110A) mutant. I tested both of these using the sedimentation assay with three biological replicates. The result (**Figure 18**) showed again that the bacterial cells were still clumping together. Even in classical autotransporter Ag43 not only one but a number of point mutations were required to prevent autoaggregation (Heras *et al.*, 2014).

(Hoiczyk *et al.*, 2000) suggested that deletion of the YeYadA head domain abolished the autoggregation. We also tested the effect of deleting of the head domain of YeYadA and also that of EibD, but this did not abrogate the autoaggregation in either case. That might be due to experimental setup was different with what did, and also may be the coiled coil that splays apart they might stick to each other and make the bacterial cell aggregated.
It might be interesting to observe the behaviour of these constructs in the microscope, but because lake of time it was not possible.
As the deletion of YLH domain did not abrogate the autoaggregation for EibD, we deleted also the N-terminal domain of EibD, which has an unknown function.
We also wanted to know whether the uptake region of YpYadA is implicated in the autoaggregation property for YpYadA, in addition to the fibronectin binding function (Heise & Dersch, 2006). The sedimentation assay was performed with one biological replicate. The results showed that the bacterial cells were still autoaggregating and settled down. Therefore, I concluded that the uptake region is not central to autoaggregation of YpYadA.

4.4 Biological implications of co-aggregation

What happens when two bacteria expressing TAAs co-aggregate? Do they generate mixed biofilms, do they cooperate, or do they compete? Under conditions that are limiting for resources, such as in the intestine, strong competition or cooperation might happen between the different bacteria expressing TAAs used in this study.
Eib proteins have some important functions including binding to Igs, autoaggregation, and biofilm formation. Some of them have been characterized from a commensal strain, ECOR 9 (Sandt & Hill, 2000, 2001).
Uncharacterized *eib* genes have been also found in another commensal *E. coli* strain, ED1a (Touchon *et al.*, 2009). The *eib* genes are located in prophage sequence in these strains, which can be mobilized (Sandt & Hill, 2000). This suggests a possible means for horizontal transfer of these genes between different strains.
This in turn suggests a mechanism for producing new TAA variants by genetic recombination, because it might be that these commensal *E. coli* strains act as a reservoir for these putative virulence factors. The properties the Eibs have would be advantageous for commensal strains as well, suggesting the Eibs could be ‘‘commensalism factors’’ (Leo *et al.*, 2011) that promote cooperation rather than competition, which means that they might promote co-aggregation between *E. coli* commensal bacteria. One strain (ECOR9) can express several Eibs, and performing it can presumably aggregate with several different commensal strains.
As an example, of co-operation between genetically related (kin) strains, *Bacillus subtilis*, a soil bacterium, forms biofilms on plant roots. Pairs of kin strains were able to co-colonize on root surfaces and formed a mixed-strain biofilm. In contrast, inoculating roots with pairs of non-kin strains resulted in biofilms consisting primarily of one strain, suggesting competition among non-kin strains (Stefanic *et al.*, 2015).

Some studies showed that the *Propionibacterium* strains showed both autoaggregation and co-aggregation with the *E. coli* strain ATTC 11229, and the binding was depending on strain and incubation conditions (Piwat *et al.*, 2015). (Re *et al.*, 2001) found that the autoaggregation of probiotic strains appeared to be necessary for adhesion to intestinal epithelial cells, and coaggregation abilities may form a barrier that prevents colonization by pathogenic microorganisms.

In contrast, competition might happen between the *Yersiniae*, because they have very similar life styles in terms of infection pathways. Both pathogens would probably be in competition with each other for obtaining resources if they infect at the same time; therefore, it makes sense for YeYadA and YpYadA to exclude each other. They probably do not form biofilm-type structures early in the infection, so the mixed biofilm might not be an issue there. Autoaggregation and coaggregation may be involved in the co-operation and competition between different bacterial strains, thus playing a role in the ecology of these organisms.

5 Conclusions and future perspectives

5.1 Conclusions

In this study, all the constructs were made successfully using the plasmid pACYCDuet -1 in a way that it can carry two different proteins (a fluorescent protein and a TAA) that can be co-expressed together. Using a cytoplasmic fluorescent marker does not affect the function of the main protein of interest, whose aggregation properties we want to characterize. The plasmid backbones in this study could thus be useful for other co-aggregation experiments using other proteins.
The results showed for the first time that some TAAs co-aggregate together and the co-aggregation was depending on sequence similarity. However, we found two instances where different TAAs excluded each other.
Some of the TAA pairs with interesting interactions were tested for biofilm formation. The results showed that the co-aggregated TAAs formed mixed biofilm while one pair of excluding TAAs segregated into microdomain within the biofilm.
None of the mutated samples made in this study (point mutants and domain deletions) abrogated the autoaggregation property of the TAAs.
Finally, since all TAAs share a common N-terminal head- stalk-membrane anchor–C-terminal architecture (Linke *et al.*, 2006), it might be that TAAs containing a particular domain such as YLH domain mediate co-aggregation or the formation of mixed biofilms.

5.2 Future perspectives

In this study, we established that different TAAs can co-aggregate and form mixed biofilm. However, there are still some technical issues should be improved and further experiments should be done for deeper understanding.
So, the first future perspective will be the improvement of the script to be able to distinguish each single bacterium. for example, the EibD-sfGFP interaction gives a high association index, which we do not think is correct because sedimentation assay showed that the sfGFP is not clumping and did not settle down or behaviour as TAA and was obvious in the graph. Furthermore, the micrograph shows that sfGFP-expressing cells are single, but some are trapped within clumps of EibD-expressing cells. These results demonstrate that the basis for calculating the association index can be improved.

The distance between the colocalized spots (red and green) that resemble the red and green bacteria, should be increased form 1.15 µm to 1.5 µm in order to include all the bacteria around that area for 3D images of biofilm analysed by Imaris.

Another possible experiment is the abrogation of the autoaggregation properties by deleting the YLH + neck domains for YeYadA construct and EibD. The current constructs have the neck and they are still aggregating. So, by also deleting the neck domain along with the YLH we might be able to abrogate the autoaggregation. This would establish the role of the neck in autoaggregation.

I did not have time to examine the mutagenesis constructs made in this study under CSLM. However, it could be interesting to do so, to see for example whether the deletion of the

uptake region in YpYadA will have any effect on the phenotype, i.e. the formation of sausage-like aggregates.

On a broader scale, co-aggregation properties of TAAs might be used for preventing or treating diseases. For example, probiotic strains have been the subject of many papers, because these non-pathogenic live organisms can provide a diverse benefit for health of the host (Kang & Im, 2015).
Therefore, one possible application of TAA-mediated co-aggregation might be to create probiotic strains. Such strain may co-aggregate with pathogenic bacteria expressing TAAs thus forming a barrier that prevents colonization.

Since the autoaggregation mediated by the TAAs used in this study is suggested to be an important virulence factor for colonization and causing infections(Trunk *et al.*, 2018), it might be possible to design a drug that can abolish this property and treat diseases caused by these bacteria.

References

Abdel-Nour, Duncan, Prashar, Rao, Ginevra, Jarraud, et al. (2014). The *Legionella pneumophila* collagen-like protein mediates sedimentation, autoaggregation, and pathogen-phagocyte interactions. Appl Environ Microbiol, 80(4), 1441-1454.

Abramoff, Magalhães, & Ram. (2003). Image Processing with ImageJ (Vol. 11).

Ai, Baird, Shen, Davidson, & Campbell. (2014). Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. Nature Protocols, 9, 910.

Alva, Nam, Söding, & Lupas. (2016). The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Research, 44(Web Server issue), W410-W415.

Arenas, Cano, Nijland, Dongen, Rutten, Ende, et al. (2015). The meningococcal autotransporter AutA is implicated in autoaggregation and biofilm formation. Environmental Microbiology, 17(4), 1321-1337.

Ausmees, Jacobsson, & Lindberg. (2001). A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to a family of Rhizobium-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. trifolii. Microbiology, 147(Pt 3), 549-559.

Balligand, Laroche, & Cornelis. (1985). Genetic analysis of virulence plasmid from a serogroup 9 Yersinia enterocolitica strain: role of outer membrane protein P1 in resistance to human serum and autoagglutination. Infect Immun, 48(3), 782-786.

Bardiau, Gregoire, Muylaert, Nahayo, Duprez, Mainil, et al. (2010). Enteropathogenic (EPEC), enterohaemorragic (EHEC) and verotoxigenic (VTEC) *Escherichia coli* in wild cervids. J Appl Microbiol, 109(6), 2214-2222.

Benz, & Schmidt. (1992). AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesized via a precursor molecule. Mol Microbiol, 6(11), 1539-1546.

Bernstein. (2015). Looks can be deceiving: recent insights into the mechanism of protein secretion by the autotransporter pathway. Mol Microbiol, 97(2), 205-215.

Biedzka-Sarek, Salmenlinna, Gruber, Lupas, Meri, & Skurnik. (2008). Functional mapping of YadA- and Ail-mediated binding of human factor H to *Yersinia enterocolitica* serotype O:3. Infect Immun, 76(11), 5016-5027.

Bolin, Norlander, & Wolf-Watz. (1982). Temperature-inducible outer membrane protein of *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* is associated with the virulence plasmid. Infect Immun, 37(2), 506-512.

Bos, van der Mei, & Busscher. (1999). Physico-chemistry of initial microbial adhesive interactions--its mechanisms and methods for study. FEMS Microbiol Rev, 23(2), 179-230.

Buswell, Herlihy, Marsh, Keevil, & Leach. (1997). Coaggregation amongst aquatic biofilm bacteria. Journal of Applied Microbiology, 83(4), 477-484.

Capecchi, Adu-Bobie, Di Marcello, Ciucchi, Masignani, Taddei, et al. (2005). *Neisseria meningitidis* NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. Molecular Microbiology, 55(3), 687-698.

Chain, Carniel, Larimer, Lamerdin, Stoutland, Regala, et al. (2004). Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. Proceedings of the National Academy of Sciences of the United States of America, 101(38), 13826-13831.

Chang, & Cohen. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. Journal of Bacteriology, 134(3), 1141-1156.

Charbonneau, Janvore, & Mourez. (2009). Autoprocessing of the Escherichia coli AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. J Biol Chem, 284(25), 17340-17351.

Chatzou, Magis, Chang, Kemena, Bussotti, Erb, et al. (2016). Multiple sequence alignment modeling: methods and applications. Brief Bioinform, 17(6), 1009-1023.

Cheng, Meng, Wang, Chen, & Li. (2014). Isolation and Characterization of Broad Spectrum Coaggregating Bacteria from Different Water Systems for Potential Use in Bioaugmentation. PLoS One, 9(4), e94220.

Chiang, Taylor, Koomey, & Mekalanos. (1995). Single amino acid substitutions in the N-terminus of *Vibrio cholerae* TcpA affect colonization, autoagglutination, and serum resistance. Mol Microbiol, 17(6), 1133-1142.

Clantin, Delattre, Rucktooa, Saint, Méli, Locht, et al. (2007). Structure of the Membrane Protein FhaC: A Member of the Omp85-TpsB Transporter Superfamily. Science, 317(5840), 957-961.

Colley, Dederer, Carnell, Kjelleberg, Rice, & Klebensberger. (2016). SiaA/D Interconnects c-di-GMP and RsmA Signaling to Coordinate Cellular Aggregation of *Pseudomonas aeruginosa* in Response to Environmental Conditions. Frontiers in Microbiology, 7, 179.

Collinson, Doig, Doran, Clouthier, Trust, & Kay. (1993). Thin, aggregative fimbriae mediate binding of Salmonella enteritidis to fibronectin. J Bacteriol, 175(1), 12-18.

Conners, Hill, Borodina, Agnew, Daniell, Burton, et al. (2008). The *Moraxella* adhesin UspA1 binds to its human CEACAM1 receptor by a deformable trimeric coiled-coil. The EMBO Journal, 27(12), 1779-1789.

Cotter, Surana, & St. Geme. Trimeric autotransporters: a distinct subfamily of autotransporter proteins. Trends in Microbiology, 13(5), 199-205.

Cover , & Aber (1989). *Yersinia Enterocolitica*. New England Journal of Medicine, 321(1), 16-24.

Dje N'Guessan, Vigelahn, Vigelahn, Bachmann, Zabel, Opitz, et al. (2007). The UspA1 Protein of *Moraxella catarrhalis* Induces CEACAM-1–Dependent Apoptosis in Alveolar Epithelial Cells. The Journal of Infectious Diseases, 195(11), 1651-1660.

Donlan. (2001). Biofilms and device-associated infections. Emerging Infectious Diseases, 7(2), 277-281.

Dunne. (2002). Bacterial Adhesion: Seen Any Good Biofilms Lately? Clinical Microbiology Reviews, 15(2), 155-166.

Dvorak, Chrast, Nikel, Fedr, Soucek, Sedlackova, et al. (2015). Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. Microbial Cell Factories, 14(1), 201.

El Tahir, & Skurnik. (2001). YadA, the multifaceted Yersinia adhesin. International Journal of Medical Microbiology, 291(3), 209-218.

El-Kirat-Chatel, Mil-Homens, Beaussart, Fialho, & Dufrene. (2013). Single-molecule atomic force microscopy unravels the binding mechanism of a *Burkholderia cenocepacia* trimeric autotransporter adhesin. Mol Microbiol, 89(4), 649-659.

Elliott, Wilson, Buckley, & Spratt. (2006). Aggregative behavior of bacteria isolated from canine dental plaque. Appl Environ Microbiol, 72(8), 5211-5217.

Emody, Heesemann, Wolf-Watz, Skurnik, Kapperud, O'Toole, et al. (1989). Binding to collagen by *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*: evidence for yopA-mediated and chromosomally encoded mechanisms. J Bacteriol, 171(12), 6674-6679.

Fan, Chauhan, Udatha, Leo, & Linke. (2016). Type V Secretion Systems in Bacteria. Microbiol Spectr, 4(1).

Felek, Lawrenz, & Krukonis. (2008). The *Yersinia pestis* autotransporter YapC mediates host cell binding, autoaggregation and biofilm formation. Microbiology, 154(Pt 6), 1802-1812.

Formosa-Dague, Feuillie, Beaussart, Derclaye, Kucharikova, Lasa, et al. (2016). Sticky Matrix: Adhesion Mechanism of the Staphylococcal Polysaccharide Intercellular Adhesin. ACS Nano, 10(3), 3443-3452.

Frick, Morgelin, & Bjorck. (2000). Virulent aggregates of *Streptococcus pyogenes* are generated by homophilic protein-protein interactions. Mol Microbiol, 37(5), 1232-1247.

Gibbons, & Nygaard. (1970). Interbacterial aggregation of plaque bacteria. Arch Oral Biol, 15(12), 1397-1400.

Gibson, Young, Chuang, Venter, Hutchison, & Smith. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods, 6(5), 343-345.

Glaubman, Hofmann, Bonney, Park, Thomas, Kokona, et al. (2016). Self-association motifs in the enteroaggregative *Escherichia coli* heat-resistant agglutinin 1. Microbiology, 162(7), 1091-1102.

Green, & Mecsas. (2016). Bacterial Secretion Systems – An overview. Microbiology spectrum, 4(1), 10.1128/microbiolspec.VMBF-0012-2015.

Grijpstra, Arenas, Rutten, & Tommassen. (2013). Autotransporter secretion: varying on a theme. Research in Microbiology, 164(6), 562-582.

Grin, Hartmann, Sauer, Hernandez Alvarez, Schutz, Wagner, et al. (2014). A trimeric lipoprotein assists in trimeric autotransporter biogenesis in enterobacteria. J Biol Chem, 289(11), 7388-7398.

Guérin, Bigot, Schneider, Buchanan, & Jacob-Dubuisson. (2017). Two-Partner Secretion: Combining Efficiency and Simplicity in the Secretion of Large Proteins for Bacteria-Host and Bacteria-Bacteria Interactions. Frontiers in Cellular and Infection Microbiology, 7(148).

Guerry. (2007). Campylobacter flagella: not just for motility. Trends Microbiol, 15(10), 456-461.

Gupta, Sarkar, Das, Bhattacharjee, & Tribedi. (2016). Biofilm, pathogenesis and prevention--a journey to break the wall: a review. Arch Microbiol, 198(1), 1-15.

Hall-Stoodley, & Stoodley. (2009). Evolving concepts in biofilm infections. Cell Microbiol, 11(7), 1034-1043.

Heise, & Dersch. (2006). Identification of a domain in *Yersinia* virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. Proc Natl Acad Sci U S A, 103(9), 3375-3380.

Henderson, Navarro-Garcia, Desvaux, Fernandez, & Ala'Aldeen. (2004). Type V protein secretion pathway: the autotransporter story. Microbiol Mol Biol Rev, 68(4), 692-744.

Henderson, Ward John, & Ready. (2010). *Aggregatibacter* (Actinobacillus) *actinomycetemcomitans*: a triple A* periodontopathogen? Periodontology 2000, 54(1), 78-105.

Heras, Totsika, Peters, Paxman, Gee, Jarrott, et al. (2014). The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. Proc Natl Acad Sci U S A, 111(1), 457-462.

Hevia, Martinez, Ladero, Alvarez, Margolles, & Sanchez. (2013). An extracellular Serine/Threonine-rich protein from *Lactobacillus plantarum* NCIMB 8826 is a

novel aggregation-promoting factor with affinity to mucin. Appl Environ Microbiol, 79(19), 6059-6066.

Hoiczyk, Roggenkamp, Reichenbecher, Lupas, & Heesemann. (2000). Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. The EMBO Journal, 19(22), 5989-5999

Iadanza, Higgins, Schiffrin, Calabrese, Brockwell, Ashcroft, et al. (2016). Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. Nature communications, 7, 12865.

Ishikawa, Nakatani, & Hori. (2012). AtaA, a new member of the trimeric autotransporter adhesins from *Acinetobacter sp*. Tol 5 mediating high adhesiveness to various abiotic surfaces. PLoS One, 7(11), e48830.

Jain, & Goldberg. (2007). Requirement for YaeT in the Outer Membrane Assembly of Autotransporter Proteins. Journal of Bacteriology, 189(14), 5393-5398.

Kaiser, Riess, Wagner, Linke, Lupas, Schwarz, et al. (2008). The head of *Bartonella* adhesin A is crucial for host cell interaction of Bartonella henselae. Cell Microbiol, 10(11), 2223-2234

Kang, & Im. (2015). Probiotics as an Immune Modulator. Journal of Nutritional Science and Vitaminology, 61(Supplement), S103-S105.

Kang'ethe, & Bernstein. (2013). Charge-dependent secretion of an intrinsically disordered protein via the autotransporter pathway. Proc Natl Acad Sci U S A, 110(45), E4246-4255.

Karatan, & Watnick. (2009). Signals, regulatory networks, and materials that build and break bacterial biofilms. Microbiol Mol Biol Rev, 73(2), 310-347.

Kjærgaard, Schembri, Hasman, & Klemm. (2000). Antigen 43 from *Escherichia coli* induces inter-and intraspecies cell aggregation and changes in colony morphology of *Pseudomonas fluorescens*. Journal of bacteriology, 182(17), 4789-4796.

Klebensberger, Rui, Fritz, Schink, & Philipp. (2006). Cell aggregation of *Pseudomonas aeruginosa* strain PAO1 as an energy-dependent stress response during growth with sodium dodecyl sulfate. Arch Microbiol, 185(6), 417-427.

Klemm, & Schembri. (2000). Bacterial adhesins: function and structure. International Journal of Medical Microbiology, 290(1), 27-35.

Klemm, Vejborg, & Sherlock. (2006). Self-associating autotransporters, SAATs: Functional and structural similarities. International Journal of Medical Microbiology, 296(4), 187-195.

Kolenbrander, Andersen, Blehert, Egland, Foster, & Palmer. (2002). Communication among Oral Bacteria. Microbiology and Molecular Biology Reviews, 66(3), 486-505.

Kolenbrander, Andersen, & Holdeman. (1985). Coaggregation of oral Bacteroides species with other bacteria: central role in coaggregation bridges and competitions. Infection and Immunity, 48(3), 741-746.

Kolodziejek, Schnider, Rohde, Wojtowicz, Bohach, Minnich, et al. (2010). Outer membrane protein X (Ail) contributes to Yersinia pestis virulence in pneumonic plague and its activity is dependent on the lipopolysaccharide core length. Infect Immun, 78(12), 5233-5243.

Koretke, Szczesny, Gruber, & Lupas. (2006). Model structure of the prototypical non-fimbrial adhesin YadA of Yersinia enterocolitica. J Struct Biol, 155(2), 154-161.

Kos, Šušković, Vuković, Šimpraga, Frece, & Matošić. (2003). Adhesion and aggregation ability of probiotic strain Lactobacillus acidophilus M92. Journal of Applied Microbiology, 94(6), 981-987.

Kragh, Hutchison, Melaugh, Rodesney, Roberts, Irie, et al. (2016). Role of Multicellular Aggregates in Biofilm Formation. MBio, 7(2), e00237.

Kuroda, Ito, Tanaka, Yao, Matoba, Saito, et al. (2008). Staphylococcus aureus surface protein SasG contributes to intercellular autoaggregation of Staphylococcus aureus. Biochem Biophys Res Commun, 377(4), 1102-1106.

L., Z., Bruce, Phillip, & W. (2003). The Haemophilus influenzae Hap autotransporter mediates microcolony formation and adherence to epithelial cells and extracellular matrix via binding regions in the C-terminal end of the passenger domain. Cellular Microbiology, 5(3), 175-186.

Lambris, Ricklin, & Geisbrecht. (2008). Complement evasion by human pathogens. Nature Reviews Microbiology, 6(2), 132.

Laporte, Savin, Lamourette, Devilliers, Volland, Carniel, et al. (2015). Fast and sensitive detection of enteropathogenic Yersinia by immunoassays. J Clin Microbiol, 53(1), 146-159.

Leo, Elovaara, Bihan, Pugh, Kilpinen, Raynal, et al. (2010). First analysis of a bacterial collagen-binding protein with collagen Toolkits: promiscuous binding of YadA to collagens may explain how YadA interferes with host processes. Infect Immun, 78(7), 3226-3236.

Leo, Elovaara, Brodsky, Skurnik, & Goldman. (2008). The Yersinia adhesin YadA binds to a collagenous triple-helical conformation but without sequence specificity. Protein Eng Des Sel, 21(8), 475-484.

Leo, & Goldman. (2009). The immunoglobulin-binding Eib proteins from *Escherichia coli* are receptors for IgG Fc. Mol Immunol, 46(8-9), 1860-1866.

Leo, Grin, & Linke. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. Philos Trans R Soc Lond B Biol Sci, 367(1592), 1088-1101.

Leo, Lyskowski, Hattula, Hartmann, Schwarz, Butcher, et al. (2011). The structure of E. coli IgG-binding protein D suggests a general model for bending and binding in trimeric autotransporter adhesins. Structure, 19(7), 1021-1030.

Leo, Oberhettinger, Chaubey, Schutz, Kuhner, Bertsche, et al. (2015). The Intimin periplasmic domain mediates dimerisation and binding to peptidoglycan. Mol Microbiol, 95(1), 80-100.

Leo, & Skurnik. (2011). Adhesins of human pathogens from the genus *Yersinia*. Adv Exp Med Biol, 715, 1-15.

Leyton, Rossiter, & Henderson. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. Nat Rev Microbiol, 10(3), 213-225.

Li, Xia, Tao, & Wang. (2017). Modeling Biofilms in Water Systems with New Variables: A Review. Water, 9(7).

Linke, Riess, Autenrieth, Lupas, & Kempf. (2006). Trimeric autotransporter adhesins: variable structure, common function. Trends Microbiol, 14(6), 264-270.

Liu, & Naismith. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol, 8, 91.

Lu, Iyoda, Satou, Satou, Itoh, Saitoh, et al. (2006). A New Immunoglobulin-Binding Protein, EibG, Is Responsible for the Chain-Like Adhesion Phenotype of Locus of Enterocyte Effacement-Negative, Shiga Toxin-Producing *Escherichia coli*. Infection and Immunity, 74(10), 5747-5755.

MacKichan, Gerns, Chen, Zhang, & Koehler. (2008). A SacB mutagenesis strategy reveals that the Bartonella quintana variably expressed outer membrane proteins are required for bloodstream infection of the host. Infect Immun, 76(2), 788-795.

Malik, Sakamoto, Hanazaki, Osawa, Suzuki, Tochigi, et al. (2003). Coaggregation among Nonflocculating Bacteria Isolated from Activated Sludge. Applied and Environmental Microbiology, 69(10), 6056-6063.

Malito, Biancucci, Faleri, Ferlenghi, Scarselli, Maruggi, et al. (2014). Structure of the meningococcal vaccine antigen NadA and epitope mapping of a bactericidal antibody. Proc Natl Acad Sci U S A, 111(48), 17128-17133.

Martínez-Gil, Yousef-Coronado, & Espinosa-Urgel. (2010). LapF, the second largest *Pseudomonas putida* protein, contributes to plant root colonization and determines biofilm architecture. Molecular Microbiology, 77(3), 549-561.

McCabe, Ricci, Adetunji, & Silhavy. (2017). Conformational Changes That Coordinate the Activity of BamA and BamD Allowing β-Barrel Assembly. Journal of Bacteriology, 199(20), e00373-00317.

Melaugh, Hutchison, Kragh, Irie, Roberts, Bjarnsholt, et al. (2016). Shaping the Growth Behaviour of Biofilms Initiated from Bacterial Aggregates. PLoS One, 11(3), e0149683.

Merkel, Ohder, Bielaszewska, Zhang, Fruth, Menge, et al. (2010). Distribution and phylogeny of immunoglobulin-binding protein G in Shiga toxin-producing *Escherichia coli* and its association with adherence phenotypes. Infect Immun, 78(8), 3625-3636.

Mikula, Leo, Lyskowski, Kedracka-Krok, Pirog, & Goldman. (2012). The translocation domain in trimeric autotransporter adhesins is necessary and sufficient for trimerization and autotransportation. J Bacteriol, 194(4), 827-838.

Milbredt, & Waldminghaus. (2017). BiFCROS: A Low-Background Fluorescence Repressor Operator System for Labeling of Genomic Loci. G3 (Bethesda), 7(6), 1969-1977.

Mintz. (2004). Identification of an extracellular matrix protein adhesin, EmaA, which mediates the adhesion of *Actinobacillus actinomycetemcomitans* to collagen. Microbiology, 150(Pt 8), 2677-2688.

Muhlenkamp, Hallstrom, Autenrieth, Bohn, Linke, Rinker, et al. (2017). Vitronectin Binds to a Specific Stretch within the Head Region of Yersinia Adhesin A and Thereby Modulates *Yersinia enterocolitica* Host Interaction. J Innate Immun, 9(1), 33-51.

Muller, Kaiser, Linke, Schwarz, Riess, Schafer, et al. (2011). Trimeric autotransporter adhesin-dependent adherence of *Bartonella henselae*, *Bartonella quintana*, and *Yersinia enterocolitica* to matrix components and endothelial cells under static and dynamic flow conditions. Infect Immun, 79(7), 2544-2553.

Mühlenkamp, Oberhettinger, Leo, Linke, & Schütz. (2015). Yersinia adhesin A (YadA) – Beauty & beast. International Journal of Medical Microbiology, 305(2), 252-258.

Noinaj, Kuszak, Gumbart, Lukacik, Chang, Easley, et al. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. Nature, 501, 385.

Nummelin, Merckel, Leo, Lankinen, Skurnik, & Goldman. (2004). The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. Embo j, 23(4), 701-711.

O'Toole, & Kolter. (1998). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. Mol Microbiol, 30(2), 295-304.

Ojanen-Reuhs, Kalkkinen, Westerlund-Wikstrom, van Doorn, Haahtela, Nurmiaho-Lassila, et al. (1997). Characterization of the fimA gene encoding bundle-forming fimbriae of the plant pathogen Xanthomonas campestris pv. vesicatoria. J Bacteriol, 179(4), 1280-1290.

Park, Wolfgang, Van Putten, Dorward, Hayes, & Koomey. (2001). Structural alterations in a type IV pilus subunit protein result in concurrent defects in multicellular behaviour and adherence to host tissue. Molecular Microbiology, 42(2), 293-307.

Parsek, & Singh. (2003). Bacterial biofilms: an emerging link to disease pathogenesis. Annu Rev Microbiol, 57, 677-701.

Pearson, Lafontaine, Wagner, Geme, & Hansen. (2002). A hag mutant of *Moraxella catarrhalis* strain O35E is deficient in hemagglutination, autoagglutination, and immunoglobulin D-binding activities. Infection and Immunity, 70(8), 4523-4533.

Pepe, Wachtel, Wagar, & Miller. (1995). Pathogenesis of defined invasion mutants of *Yersinia enterocolitica* in a BALB/c mouse model of infection. Infection and Immunity, 63(12), 4837-4848.

Pérez-Ortega, Rodríguez, Ribes, Tommassen, & Arenas. (2017). Interstrain Cooperation in Meningococcal Biofilms: Role of Autotransporters NalP and AutA. Frontiers in Microbiology, 8, 434.

Peterson, Tian, Ieva, Dautin, & Bernstein. (2010). Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. Proc Natl Acad Sci U S A, 107(41), 17739-17744.

Piwat, Sophatha, & Teanpaisan. (2015). An assessment of adhesion, aggregation and surface charges of *Lactobacillus* strains derived from the human oral cavity. Letters in Applied Microbiology, 61(1), 98-105.

Podladchikova, Rykova, Antonenka, & Rakin. (2012). *Yersinia pestis* autoagglutination is mediated by HCP-like protein and siderophore Yersiniachelin (Ych). Adv Exp Med Biol, 954, 289-292.

Raghunathan, Wells, Morris, Shaw, Bobat, Peters, et al. (2011). SadA, a trimeric autotransporter from *Salmonella enterica* serovar Typhimurium, can promote biofilm formation and provides limited protection against infection. Infect Immun, 79(11), 4342-4352.

Re, Sgorbati, Miglioli, & Palenzona. (2001). Adhesion, autoaggregation and hydrophobicity of 13 strains of Bifidobacterium longum (Vol. 31).

Reid, McGroarty, Angotti, & Cook. (1988). Lactobacillus inhibitor production against *Escherichia coli* and coaggregation ability with uropathogens. Can J Microbiol, 34(3), 344-351.

Renn, Junker, Besingi, Braselmann, & Clark. (2012). ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. Chemistry & biology, 19(2), 287-296.

Reuter, Connor, Barquist, Walker, Feltwell, Harris, et al. (2014). Parallel independent evolution of pathogenicity within the genus *Yersinia*. Proceedings of the National Academy of Sciences, 111(18), 6768-6773.

Rickard, Gilbert, & Handley. (2004). Influence of growth environment on coaggregation between freshwater biofilm bacteria. J Appl Microbiol, 96(6), 1367-1373.

Rickard, Gilbert, High, Kolenbrander, & Handley. (2003). Bacterial coaggregation: an integral process in the development of multi-species biofilms. Trends Microbiol, 11(2), 94-100.

Rickard, McBain, Ledder, Handley, & Gilbert. (2003). Coaggregation between freshwater bacteria within biofilm and planktonic communities. FEMS Microbiology Letters, 220(1), 133-140.

Riesbeck, Tan, & Forsgren. (2006). MID and UspA1/A2 of the human respiratory pathogen *Moraxella catarrhalis*, and interactions with the human host as basis for vaccine development. ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-, 53(3), 445.

Roberts, Rudolf, Meyer, Pellaux, Whitehead, Panke, et al. (2016). Identification and Characterisation of a pH-stable GFP. Sci Rep, 6, 28166.

Roggenkamp, Ackermann, Jacobi, Truelzsch, Hoffmann, & Heesemann. (2003). Molecular Analysis of Transport and Oligomerization of the *Yersinia enterocolitica* Adhesin YadA. Journal of Bacteriology, 185(13), 3735-3744.

Roggenkamp, Neuberger, Flügel, Schmoll, & Heesemann. (1995). Substitution of two histidine residues in YadA protein of *Yersinia enterocolitica* abrogates collagen binding, cell adherence and mouse virulence. Molecular Microbiology, 16(6), 1207-1219.

Rohde, Burdelski, Bartscht, Hussain, Buck, Horstkotte, et al. (2005). Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. Mol Microbiol, 55(6), 1883-1895.

Roman-Hernandez, Peterson, & Bernstein. (2014). Reconstitution of bacterial autotransporter assembly using purified components. Elife, 3, e04234.

Rosano, & Ceccarelli. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. Frontiers in Microbiology, 5, 172.

Rosen, Genzler, & Sela. (2008). Coaggregation of Treponema denticola with *Porphyromonas gingivalis* and *Fusobacterium nucleatum* is mediated by the major outer sheath protein of Treponema denticola. FEMS Microbiol Lett, 289(1), 59-66.

Rouviere, & Gross. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. Genes Dev, 10(24), 3170-3182.

Sandt, & Hill. (2000). Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of *Escherichia coli*. Infect Immun, 68(4), 2205-2214.

Sandt, & Hill. (2001). Nonimmune binding of human immunoglobulin A (IgA) and IgG Fc by distinct sequence segments of the EibF cell surface protein of *Escherichia coli*. Infect Immun, 69(12), 7293-7303.

Sandt, Wang, Wilson, & Hill. (1997). *Escherichia coli* strains with nonimmune immunoglobulin-binding activity. Infection and Immunity, 65(11), 4572-4579.

Sanger, & Coulson. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol, 94(3), 441-448.

Sasaki, Ishikawa, Terayama, Asano, Kawamoto, Ishibashi, et al. (2016). Identification of a virulence determinant that is conserved in the Jawetz and Heyl biotypes of [*Pasteurella*] pneumotropica. Pathog Dis, 74(6).

Sathiamoorthy, & Shin. (2012). Boundaries of the Origin of Replication: Creation of a pET-28a-Derived Vector with p15A Copy Control Allowing Compatible Coexistence with pET Vectors (Vol. 7).

Satpathy, Sen, Pattanaik, & Raut. (2016). Review on bacterial biofilm: An universal cause of contamination. Biocatalysis and Agricultural Biotechnology, 7, 56-66.

Schafer, Beck, & Muller. (1999). Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. J Biol Chem, 274(35), 24567-24574.

Schembri, Christiansen, & Klemm. (2001). FimH-mediated autoaggregation of *Escherichia coli*. Molecular Microbiology, 41(6), 1419-1430.

Schembri, & Klemm. (2001). Coordinate gene regulation by fimbriae-induced signal transduction. The EMBO Journal, 20(12), 3074-3081.

Schiffrin, Calabrese, Higgins, Humes, Ashcroft, Kalli, et al. (2017). Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. Journal of Molecular Biology, 429(23), 3776-3792.

Schulze-Koops, Burkhardt, Heesemann, Von Der Mark, & Emmrich. (1992). Plasmid-encoded outer membrane protein YadA mediates specific binding of enteropathogenic yersiniae to various types of collagen. Infection and Immunity, 60(6), 2153-2159.

Serruto, Spadafina, Scarselli, Bambini, Comanducci, Hohle, et al. (2009). HadA is an atypical new multifunctional trimeric coiled-coil adhesin of *Haemophilus influenzae* biogroup aegyptius, which promotes entry into host cells. Cell Microbiol, 11(7), 1044-1063.

Sheets, & St. Geme. (2011). Adhesive Activity of the *Haemophilus* Cryptic Genospecies Cha Autotransporter Is Modulated by Variation in Tandem Peptide Repeats. Journal of Bacteriology, 193(2), 329-339.

Sherlock, Schembri, Reisner, & Klemm. (2004). Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. J Bacteriol, 186(23), 8058-8065.

Sherlock, Vejborg, & Klemm. (2005). The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. Infect Immun, 73(4), 1954-1963.

Shimomura, Johnson Frank, & Saiga. (2005). Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the *Luminous Hydromedusan*, Aequorea. Journal of Cellular and Comparative Physiology, 59(3), 223-239.

Sikdar, Peterson, Anderson, & Bernstein. (2017). Folding of a bacterial integral outer membrane protein is initiated in the periplasm. Nature communications, 8, 1309.

Simoes, Simoes, & Vieira. (2008). Intergeneric coaggregation among drinking water bacteria: evidence of a role for *Acinetobacter calcoaceticus* as a bridging bacterium. Appl Environ Microbiol, 74(4), 1259-1263.

Singh, Alvarado-Kristensson, Johansson, Hallgren, Westergren-Thorsson, Morgelin, et al. (2016). The Respiratory Pathogen *Moraxella catarrhalis* Targets Collagen for Maximal Adherence to Host Tissues. MBio, 7(2), e00066.

Skurnik, Bolin, Heikkinen, Piha, & Wolf-Watz. (1984). Virulence plasmid-associated autoagglutination in *Yersinia* spp. J Bacteriol, 158(3), 1033-1036.

Skurnik, & Toivanen. (1992). LcrF is the temperature-regulated activator of the yadA gene of *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. Journal of Bacteriology, 174(6), 2047-2051.

Skurnik, & Wolf-Watz. (1989). Analysis of the yopA gene encoding the Yop1 virulence determinants of Yersinia spp. Mol Microbiol, 3(4), 517-529.

Sorroche, Spesia, Zorreguieta, & Giordano. (2012). A Positive Correlation between Bacterial Autoaggregation and Biofilm Formation in Native *Sinorhizobium meliloti* Isolates from Argentina. Applied and Environmental Microbiology, 78(12), 4092-4101.

Stefanic, Kraigher, Lyons, Kolter, & Mandic-Mulec. (2015). Kin discrimination between sympatric *Bacillus subtilis* isolates. Proceedings of the National Academy of Sciences, 112(45), 14042.

Stones, & Krachler. (2016). Against the tide: the role of bacterial adhesion in host colonization. Biochemical Society Transactions, 44(6), 1571-1580.

Stoodley, Sauer, Davies, & Costerton. (2002). Biofilms as complex differentiated communities. Annu Rev Microbiol, 56, 187-209.

Stothard. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques, 28(6), 1102, 1104.

Studier. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification, 41(1), 207-234.

Tamm, Tarkkanen, Korhonen, Kuusela, Toivanen, & Skurnik. (1993). Hydrophobic domains affect the collagen-binding specificity and surface polymerization as well as the virulence potential of the YadA protein of *Yersinia enterocolitica*. Mol Microbiol, 10(5), 995-1011.

Tan, & Darby. (2004). A Movable Surface: Formation of *Yersinia sp.* Biofilms on Motile Caenorhabditis elegans. Journal of Bacteriology, 186(15), 5087-5092.

Tertti, Skurnik, Vartio, & Kuusela. (1992). Adhesion protein YadA of Yersinia species mediates binding of bacteria to fibronectin. Infection and Immunity, 60(7), 3021-3024.

Thanassi. (2011). The long and the short of bacterial adhesion regulation. J Bacteriol, 193(2), 327-328.

Touchon, Hoede, Tenaillon, Barbe, Baeriswyl, Bidet, et al. (2009). Organised Genome Dynamics in the *Escherichia coli* Species Results in Highly Diverse Adaptive Paths. PLOS Genetics, 5(1), e1000344.

Travier, Guadagnini, Gouin, Dufour, Chenal-Francisque, Cossart, et al. (2013). ActA promotes *Listeria monocytogenes* aggregation, intestinal colonization and carriage. PLoS Pathog, 9(1), e1003131.

Tribedi, & Sil. (2014). Cell surface hydrophobicity: a key component in the degradation of polyethylene succinate by *Pseudomonas* sp. AKS2. J Appl Microbiol, 116(2), 295-303.

Trunk, S. Khalil, & Leo (2018). Bacterial autoaggregation. AIMS Microbiology, 4, 140-164.

Ulett, Webb, & Schembri. (2006). Antigen-43-mediated autoaggregation impairs motility in Escherichia coli. Microbiology, 152(7), 2101-2110.

Valle, Mabbett, Ulett, Toledo-Arana, Wecker, Totsika, et al. (2008). UpaG, a new member of the trimeric autotransporter family of adhesins in uropathogenic *Escherichia coli*. J Bacteriol, 190(12), 4147-4161.

Van Houdt, & Michiels. (2005). Role of bacterial cell surface structures in Escherichia coli biofilm formation. Research in Microbiology, 156(5), 626-633.

van Ulsen, Rahman, Jong, Daleke-Schermerhorn, & Luirink. (2014). Type V secretion: from biogenesis to biotechnology. Biochim Biophys Acta, 1843(8), 1592-1611.

van Zyl, Deane, & Dicks. (2015). Use of the mCherry Fluorescent Protein To Study Intestinal Colonization by Enterococcus mundtii ST4SA and *Lactobacillus plantarum* 423 in Mice. Applied and Environmental Microbiology, 81(17), 5993-6002.

Velarde, & Nataro. (2004). Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger. J Biol Chem, 279(30), 31495-31504.

Wang, Hsieh, Tan, Shien, Ou, Chen, et al. (2014). The haemagglutinin of *Avibacterium paragallinarum* is a trimeric autotransporter adhesin that confers haemagglutination, cell adherence and biofilm formation activities. Veterinary Microbiology174(3), 474-482.

Waters. (2009). Accuracy and precision in quantitative fluorescence microscopy. The Journal of Cell Biology, 185(7), 1135-1148.

Wolska, Grudniak, Rudnicka, & Markowska. (2016). Genetic control of bacterial biofilms. J Appl Genet, 57(2), 225-238.

Wu, Wu, & Kaiser. (1997). The *Myxococcus xanthus* pilT locus is required for social gliding motility although pili are still produced. Mol Microbiol, 23(1), 109-121.

Xiao, Zhou, Sun, Feng, Du, Gao, et al. (2012). Apa is a trimeric autotransporter adhesin of Actinobacillus pleuropneumoniae responsible for autoagglutination and host cell adherence (Vol. 52).

Zhang, Chomel, Schau, Goo, Droz, Kelminson, et al. (2004). A family of variably expressed outer-membrane proteins (Vomp) mediates adhesion and autoaggregation in *Bartonella quintana*. Proc Natl Acad Sci U S A, 101(37), 13630-13635.

Zhou, Li, & Qi. (2016). Identification and characterization of a haem biosynthesis locus in Veillonella. Microbiology, 162(10), 1735-1743.

Zhou, Liu, Merritt, & Qi. (2015). A YadA-like autotransporter, Hag 1, in *Veillonella atypica* is a Multivalent Hemagglutinin Involved in Adherence to Oral Streptococci, Porphyromonas gingivalis, and Human Oral Buccal Cells. Molecular oral microbiology, 30(4), 269-279.

Zimmermann, Stephens, Nam, Rau, Kubler, Lozajic, et al. (2017). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol.

Appendix 1

1.Protein sequence for all the TAAs used in this study.

Note / when I aligned the protein sequences for all the TAAs together I did not include the protein sequence of translocation domain that is coloured in red. Signal peptides are not included in the sequences below.

1. YeYadA (GenBank: CAA32086.1)

>YeYadA

DDYDGIPNLTAVQISPNADPALGLEYPVRPPVPGAGGLNASAKGIHSIAIGATAEAAKGAAVAV
GAGSIATGVNSVAIGPLSKALGDSAVTYGAASTAQKDGVAIGARASTSDTGVAVGFNSKADAKN
SVAIGHSSHVAANHGYSIAIGDRSKTDRENSVSIGHESLNRQLTHLAAGTKDTDAVNVAQLKKE
IEKTQENTNKRSAELLANANAYADNKSSSVLGIANNYTDSKSAETLENARKEAFAQSKDVLNM
AKAHSNSVARTTLETAEEHANSVARTTLETAEEHANKKSAEALASANVYADSKSSHTLKTANS
YTDVTVSNSTKKAIRESNQYTDHKFRQLDNRLDKLDTRVDKGLASSAALNSLFQPYGVGKVNF
TAGVGGYRSSQALAIGSGYRVNENVALKAGVAYAGSSDVMYNASFNIEW

Head domain =216 amino acid
Neck domain = (217-232) =15 amino acid
Stalk domain = (233-326) =140 amino acid
Translocation domain= (326-430) =105 amino acid

2. YpYadA (GenBank: AJJ04954.1)

> YpYadA

EEPEDGNDGIPRLSAVQISPNVDPKLGVGLYPAKPILRQENPKLPPRGLEKKRARLAEAIQPQVL
GGLDARAKGLYSIAIGATAEAAKPAAVAVGSGSMATGVNSVAIGPLSKALGDSAVTYGVSSTAQ
KDGVAIGARASASDTSVAVGFNSKVDAQNSVAIGHSSHVAADHGYSIAIGDHSKTDRENSVSIG
HESLNRQLTHLAAGTEDTDAVNVAQLKKEMAETLENARKETLAQSNDVLDAAKKHSNSVAR
TTLETAEEHANKKSAETLVSAKVYADSNSSQTLKTANSYTDVTVSNSTKKATRESNQYTDHKF
SQLDNRLDKLDKRVDKGLASSAALNSLFQPYGVGKVNFTAGVGGYRSSQALAIGSGYRVNESVA
LKAGVAYAGSSNVMYNASFNIEW

YLH domain = (1-225) =225 amino acid
Uptake region = (36-64) = 33 amino acid
(Neck + stalk) domains = (226- 231) = 106
Translocation domain = (231-305) = 73 amino acid

3. EibA (GenBank: AAF63234.1)

>EibA

QSYSALNAQNGAGSIYKVYYNPDNKTAHIDWGGLGDVEKERNKPIPLLSKIDGNGNVTITSADG
STTFTVYDKEVHDFMKAAASGKTDDIKTNLLTEQNIRDLYNRVSAIQQMETNVGLDEYGNVA
VTPNEIKERVSLQRYLAWESANSTIVANELEAQKGKLDAQKGELEAQKKNLGELTTRTDKIDA

AAAATAAKVESRTLVGVSSDGTLTRAEGAKNTISVNDGLVALSGRTDRIDAAVGAIDGRVTRN
TQSIEKNSKAIAANTRTLQQHSARLDSQQRQINENHKEMKRAAAQSAALTGLFQPYSVGKFNA
SAAVGGYSDEQALAVGVGYRFNEQTAAKAGVAFSDGDASWNVGVNFEF
N-terminal domain= (1-122) = 122
(Neck + saddle +stalk) domains = (123 + 224) =102 amino acid
Translocation domain = (224 + 363) = 139 amino acid

4. EibC (GenBank: AAF63035.1)

>EibC
QEEKYTVPYAIGEGKWGNTYEVVKTGGNGNFRYEVKEKNGKKRSLFTFDSKGDVIINGSGITYT
IHDGALNDFAQTAEKKKNGQSQSHRMTDSVVRDVYNKVYSLQRTKITGFSVEDGENGKVSLGS
DAKASGEFSVAVGTGARADKKFATAVGSWAAADGKQSTALGVGAYAYANASTAAGTAAYVDG
SAIYGTAIGNYAKVDENATEGTALGAKATVTNKNSVALGANSVTTRDNEVYIGYKTGTESDKT
YGTRVLGGLSDGTRNSDAATVGQLNRKVGGVYDDVKARITVESEKQKKYTDQKTSEVNEKVE
ARTTVGVDSDGKLTRAEGATKTIAVNDGLVALSGRTDRIDYAVGAIDGRVTRNTQSIEKNSKAI
AANTRTLQQHSARLDSQQRQINENHKEMKRAAAQSAALTGLFQPYSVGKFNATAAVGGYSDQ
QALAVGVGYRFNEQTAAKAGVAFSDGDASWNVGVNFEF

N-terminal domain = (1 - 126) =126 amino acid
YLH + neck domain = (127- 175) = 48 amino acid
Stalk + saddle = (176 -341) = 165 amino acid
Translocation domain = (342 - 481) =139

5. EibD (GenBank: AAF63040.1)

>EibD
QNGTYSVLQDDSQKSGPVKYGSTYEVVKTVDNGNFRYEVKEKKNDKRTLFKFDSEGNVTVKG
KGITHTLHDPALKDFARTAEGKKNEQNGNTPPHKLTDSAVRGVYNKVYGLEKTEITGFSVEDG
ENGKVSLGSDAKASGEFSVAVGNGARATEKASTAVGSWAAADGKQSTALGVGTYAYANASTA
LGSVAFVDNTATYGTAAGNRAKVDKDATEGTALGAKATVTNKNSVALGANSVTTRDNEVYIG
YKTGTESDKTYGTRVLGGLSDGTRNSDAATVGQLNRKVGGVYDDVKARITVESEKQKKYTDQ
KTSEVNEKVEARTTVGVDSDGKLTRAEGATKTIAVNDGLVALSGRTDRIDYAVGAIDGRVTRN
TQSIEKNSKAIAANTRTLQQHSARLDSQQRQINENHKEMKRAAAQSAALTGLFQPYSVGKFNA
TAAVGGYSDQQALAVGVGYRFNEQTAAKAGVAFSDGDASWNVGVNFEF

N-terminal domain = (1-133) =133 amino acid
YLH = (160-201) =40 amino acid
Neck = (202-210) = 9 amino acid
Salk + saddle = (211-376) =165 amino acid
Translocation domain = (212-351) = 139 amino acid

2.Multiple sequence alignment for all the TAAs

1. For comparing the Eibs group

Including YadA like head + stalk

```
EibA   QSYSALNAQNGAGSIYKVYYNPDNKTAHIDWGGLGDV-----------EKERNKPIPLLSKIDGNGNVTITSADGSTTFTVYDK
EibC   ----------QEEKYTV----PYAIGEGKWGNTYEVVKTGGNGNFRYEVKEKNGKKRSLFTFDSKGDVIINGS--GITYTIHDG
EibD   -----------QNGTYSVLQDDSQKSGPVKYGSTYEVVKTVDNGNFRYEVKEKKNDKRTLFKFDSEGNVTVKGK-GITHTLHDP

EibA   EVHDFMKAAASGKT---DDIKTNLLTEQNIRDLYNRVSAIQQMETN-------------------------------------
EibC   ALNDFAQTAEKKK---NGQSQSHRMTDSVVRDVYNKVYSLQRTKITGFSVEDGENGKVSLGSDAKASGEFSVAVGTGARADKKFA
EibD   ALKDFARTAEGKKNEQNGNTPPHKLTDSAVRGVYNKVYGLEKTEITGFSVEDGENGKVSLGSDAKASGEFSVAVGNGARATEKAS

EibA   ----------------VGLD--E-------------------YG----------------------------NVAVTPNEIKER
EibC   TAVGSWAAADGKQSTALGVGAYAYANASTAAGTAAYVDGSAIYGTAIGNYAKVDENATEGTALGAKATVTNKNSVALGANSVTTR
EibD   TAVGSWAAADGKQSTALGVGTYAYANASTALGSVAFVDNTATYGTAAGNRAKVDKDATEGTALGAKATVTNKNSVALGANSVTTR

EibA   VSLQRYLAWESANST------IVA--------NELEAQKGKLDAQKGELEAQ-KKNLG-ELTTRTDKIDAAAAATAAKVESRTLV
EibC   -DNEVYIGYKTGTESDKTYGTRVLGGLSDGTRNSDAATVGQLNRKVGGVYDDVKARITVESEKQKKYTDQKTSEVNEKVEARTTV
EibD   -DNEVYIGYKTGTESDKTYGTRVLGGLSDGTRNSDAATVGQLNRKVGGVYDDVKARITVESEKQKKYTDQKTSEVNEKVEARTTV

EibA   GVSSDGTLTRAEGAKNTISVNDGLVALSGRTDRIDAAVGAIDGRVTRNTQSIEKNSKAIAANTRTL
EibC   GVDSDGKLTRAEGATKTIAVNDGLVALSGRTDRIDYAVGAIDGRVTRNTQSIEKNSKAIAANTRTL
EibD   GVDSDGKLTRAEGATKTIAVNDGLVALSGRTDRIDYAVGAIDGRVTRNTQSIEKNSKAIAANTRTL
```

2. For comparing the YadA group with each other

Including YadA like head + stalk

```
YeYadA  ----DDYDGIPNLTAVQISPNADPALGLEYP-VRPP----------------------------VPGAGGLNASAKGIHSIAIGA
YpYadA EEPEDGNDGIPRLSAVQISPNVDPKLGVGLYPAKPILRQENPKLPPRGPQGPEKKRARLAEAIQPQVLGGLDARAKGIHSIAIGA

YeYadA  TAEAAKGAAVAVGAGSIATGVNSVAIGPLSKALGDSAVTYGAASTAQKDGVAIGARASTSDTGVAVGFNSKADAKNSVAIGHSSH
YpYadA  TAEAAKPAAVAVGAGSIATGVNSVAIGPLSKALGDSAVTYGASSTAQKDGVAIGARASASDTGVAVGFNSKVDAQNSVAIGHSSH

YeYadA  VAANHGYSIAIGDRSKTDRENSVSIGHESLNRQLTHLAAGTKDTDAVNVAQLKKEIEKTQENTNKRSAELLANANAYADNKSSSV
YpYadA  VAADHGYSIAIGDLSKTDRENSVSIGHESLNRQLTHLAAGTKDNDAVNVAQLKKEMAE--------------------------

YeYadA  LGIANNYTDSKSAETLENARKEAFAQSKDVLNMAKAHSNSVARTTLETAEEHANSVARTTLETAEEHANKKSAEALASANVYADS
YpYadA  -------------TLENARKETLAQSNDVLDAAKKHSNS--------------VARTTLETAEEHANKKSAEALVSAKVYADS

YeYadA  KSSHTLKTANSYTDVTVSNSTKKAIRESNQYTDHKFRQLDNRLDKLDT
YpYadA  NSSHTLKTANSYTDVTVSSSTKKAISESNQYTDHKFSQLDNRLDKLDK
```

3.For comparing the co-aggregated TAAs (YeYadA+EibD)

Including YadA like head + stalk

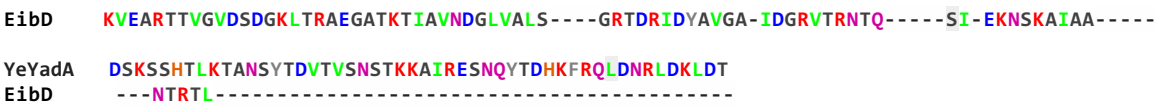
```
YeYadA  ------------------------------------------------------------------------DDYD----GI-
EibD    QNGTYSVLQDDSQKSGPVKYGSTYEVVKTVDNGNFRYEVKEKKNDKRTLFKFDSEGNVTVKGKGITHTLHDPALKDFARTAEGKK

YeYadA   -------P--NLTAVQISPNADPALGLEYPVRPPVPGAGGLNASAKGIHSIAIGATAEAAKGAAVAVGAGSIATGVNSVAIGPLS
EibD    NEQNGNTPPHKLTDSAVRGVYNKVYGLEKTEI------TGFSVEDGENGKVSLGSDAKASGEFSVAVGNGARATEKASTAVGSWA

YeYadA  -------KALGDSAVTYGAASTAQK----------DGVAIGARAS---TSDTGVAVGFNSKADAKNSVAIGHSSHVAANHGYSIA
EibD    AADGKQSTALGVGTYAYANASTALGSVAFVDNTATYGTAAGNRAKVDKDATEGTALGAKATVTNKNSVAL---------------

YeYadA  IGDRSKTDRENSVSIGHES--------LNRQLTHLAAGTKDTDAVNVAQLKKEIEKTQENTNKRSAELLANANAYADNKSSSVLG
EibD    -GANSVTTRDNEVYIGYKTGTESDKTYGTRVLGGLSDGTRNSDAATVGQLNRKVGGVYDDVKARITVESEKQKKYTDQKTSEVNE

YeYadA  IANNYTD--SKSAETLENA--RKEAFAQSKDVLNMAKAHSNSVARTTLETAEEHANSVARTTLETAEEHANKKSAEALASANVYA
```



Appendix 2 constructs and primers

Table 1. List of primers used for insert cloning in this study with purpose

Name	sequence	purpose
Duet-sfGFP	Fwd: TAA GGA GAT ATA CCA TAT GTC AAA AGG TGA AGA ATT ATT TA Rev: TGT TCG ACT TAA GCA TTA TTT ATA TAA TTC ATC CAT ACC ATG TG	For cloning into 1st MCS of Duet vectors.
Duet-mCherry	Fwd: GAA GGA GAT ATA CAT ATG GTG AGC AAG GGC GAG GAG G Rev: AGC AGC CTA GGT TAA TTA CTT GTA CAG CTC GTC CAT GCC	For cloning into 2nd MCS of Duet vectors.
Duet-EibD	Fwd: GAA GGA GAT ATA CAT ATG AGT AAA AAG TTT ACA ATG ACA CTC CT Rev: TGT TCG ACT TAA GCA TTA AAA CTC GAAGTTCACACCAAC	For amplifying EibD signal peptide to clone into 2nd MCS of Duet vectors; for cloning EibC from pET-Duets.
Duet-EibC	Is the (EiDf) Fwd: GAA GGA GAT ATA CAT ATG AGT AAA AAG TTT ACA ATG ACA CTC CT Rev: AGC AGC CTA GGT TAA TTA AAA CTC GAA GTT CAC ACC AAC	For cloning EibC into 2nd MCS of Duet vectors.
Duet-YadA	Is the pelB in the MCS2 duet Fwd: GAA GGA GAT ATA CAT ATG AAA TAC CTG CTG CCG ACC Rev: AGC AGC CTA GGT TAA TTA CCA CTC GAT ATT AAA TGA TGC ATT G	For cloning the YeYadA into 2nd MCS of Duet vector.
Duet-EibA	Is the same EibDf Fwd: GAA GGA GAT ATA CAT ATG AGT AAA AAG TTT ACA ATG ACA CTC CT Is the same EibCr Rev: AGC AGC CTA GGT TAA TTA AAA CTC GAA GTT CAC ACC AAC	For cloning the EibA into 2nd MCS of Duet vector.
Duet-YPIII	Fwd: TAA GGA GAT ATA CCA TGA CTA AAG ATT TTA AGA TCA GTG TCT CTG Rev: TGT TCG ACT TAA GCA TTA CCAC TCG ATA TTA AAT GAT GCG TT	For cloning YpYadA into the 1st MCS of Duet vector.
DuetMCS1-mCherry	Fwd: TAA GGA GAT ATA CCA TGG TGA GCA AGG GCG AGG AGG Rev: TGT TCG ACT TAA GCA TTA CTT GTA CAG CTC GTC CAT GCC	For cloning mCherry into 1st MCS of Duet vectors.YeYadA at MCS2(as a control).
DuetMCS2-sfGFP	Fwd: GAA GGA GAT ATA CAT ATG TCA AAA GGT GAA GAA TTA TTT A REV: AGC AGC CTA GGT TAA TTA TTT ATA TAA TTC ATC CAT ACC ATG TG	EibD at MCS1 and sfGFP for cloning into 2nd MCS of Duet vectors (as a control).

Table 2. Sequencing primers used in the study

Name	Primer sequence (5´–3´)
pACYCDeut -1(MCS1)	Fwd: GGA TCT CGA CGC TCT CCC T Rev: GCT AGT TAT TGC TCA GCG G
pACYCDeut -1(MCS2)	Fwd: TTG TAC ACG GCC GCA TAA TC Rev: GCT AGT TAT TGC TCA GCG G

Table 3. list of primers used for the mutagenesis

Mutation	Primer sequence (5´–3´)
Delta_EibD N	Fwd: GGT AGC GAT GCG AAA GCC TCT GGT GAG TTC TCA Rev: CGC ATC GCT ACC GTG ATG GTG ATG GTG ATG ATC C
Delta_EibD YLH	Fwd: ACC CGT GTT CTT GG TGG CTT AAG CGA TGG TAC GC Rev: CCA AGA ACA CGG GT GTG ATG GTG ATG GTG ATG ATC C
YeYadA K 68 A	Fwd: GCT AGC GCT GCG GGT ATC CAT AGC ATT GCG Rev: CGC AGC GCT AGC ATT GAG CCC GCC TGC GC
YeYadA E 182 A	Fwd: ACT GAC CGA GCG AAT AGT GTA TCC ATT GGT Rev: CGC TCG GTC AGT TTT AGA ACG ATC CCC AAT
YeYadA A 165 D	Fwd: CAC GTT GCG GAC AAT CAT GGT TAT TCA ATT GCA ATT G Rev: GTC CGC AAC GTG ACT AGA GTG TCC AAT GGC AAC A
YeYadA Delta YLH	Fwd: AGC CAG GAT CCG AGC CTT AAT CGC CAA TTG ACA CA Rev: CGG ATC CTG GCT GTG GTG ATG ATG GTG ATG GC
YeYadA E 80 A	Fwd: TGC TAC TGC T GCA GCA GCG AAA GGA GCA GCA G Rev: TGC AGC AGT AGC A CCA ATC GCA ATG CTA TGG ATA C
YeYadA I 94 A	Fwd: CGC TGG TTC A GCT GCA ACA GGC GTT AAT TCT GTT G Rev: AGC TGA ACC AGC G CCC ACA GCA ACT GCT GCT C
YeYadA L 110 A	Fwd: TA AGT AAG GCA GCG GGA GAT TCG GCA GTT ACT TAT G Rev: CGC TGC CTT ACT TA A AGG ACC AAT TGC AAC AGA ATT AAC
YeYadA N 166 A	Fwd: GTT GCG GCA GCC CAT GGT TAT TCA ATT GCA ATT GGG Rev: GGC TGC CGC AAC GTG ACT AGA GTG TCC AAT GGC
YeYadA Q124 A	Fwd: GTA CCG CC GCG AAA GAT GGA GTA GCT ATC GGT Rev: CGC GGC GGT AC TAG CTG CCC CAT AAG TAA CTG
YpYadA Delta uptake	Fwd: GTA CTA GGC GCA G GCG GGC TCA ATG CTC GC Rev: CTG CGC CTA GTA C TTT TGC TGG ATA TAA TCC CAC AC

Appendix 3 Script that used for analyzing the 2D images from CSLM.

```
//@File(label="Directory with images",style="directory") dir
files=getFileList(dir);

p="[Summary]";

if (isOpen("Summary")) {
	print(p,"\\Clear");
} else {
	run("New... ", "name="+p+" type=Table");
}
print(p,"\\Headings:File	RedToGreen	GreenToRed");

for (i=0; i< files.length; i++) {
	if (endsWith(files[i],".jpg") || endsWith(files[i],".tif") || endsWith(files[i],".tiff")) {
		open(dir + File.separator + files[i]);

		skipRest = false;

		title=getTitle();
		rename('image');

		if (bitDepth() == 8 && nSlices == 3) {
			run("Stack to Images");
			selectWindow("Blue");
			close();
		} else if (bitDepth() == 8 && nSlices == 2) {
			run("Stack to Images");
		} else if (bitDepth() == 8 && nSlices == 1) {
	showMessage(title + " contains a single channel only. This script is expecting images with 2+
	channels (or RGB image).");
			skipRest = true;
		} else if (bitDepth() == 24) {
			run("Split Channels");
			selectWindow("image (blue)");
			close();
			selectWindow("image (red)");
			rename("Red");
			selectWindow("image (green)");
}

		if (!skipRest) {
			proximityOfRedToGreen=proximityTest("Green","Red");
			proximityOfGreenToRed=proximityTest("Red","Green");

			selectWindow("Red");
			close();
			selectWindow("Green");
			close();
			print(p,title + "	" + proximityOfRedToGreen + "	"	+
			proximityOfGreenToRed);
		}
	}
}
```

```
function proximityTest(ChannelA,ChannelB) {
	roiManager("reset");

	selectWindow(ChannelA);
	setAutoThreshold("Otsu dark");
	run("Create Selection");
	roiManager("Add");
	run("Enlarge...", "enlarge=15");
	roiManager("Add");
	roiManager("Select", newArray(0,1));
	roiManager("XOR");
	roiManager("Add");
	getStatistics(potentialArea);
	selectWindow(ChannelB);
	setAutoThreshold("Otsu dark");
	run("Create Selection");
	roiManager("Add");
	roiManager("Select", newArray(2,3));
	roiManager("AND");
	roiManager("Add");

	getStatistics(coveredArea);
	return (coveredArea/potentialArea);
}
```

Appendix 4
1.Buffers and solutions

All the buffers and solutions used in this study, cited in material and methods

Agarose gel 1% (0.5L)
5 g SeaKem® LE Agarose (LONZA)
500 ml 1x TAE Buffer
The solution was heated up in the microwave in order to dissolve the agarose, then kept in the oven at 60 ^{0}C until use.

Agarose gel 0.8% (0.5L)
4 g SeaKem® LE Agarose (LONZA)
500 ml 1x TAE Buffer
The solution was heated up in the microwave in order to dissolve the agarose, then kept in the oven at 60 ^{0}C until use.

Chloramphenicol stock solution 25mg/ml (50ml)
1.25 g Cm(AppliChem)
50 ml 100% ETOH Filtered through 0.42 µm filter, store at -20^{0}C. Used at the final concentration of 25µg/ml.

DNA sample buffer 6x (10ml)
20 µl 0.5 M EDTA pH 8 (AppliChem)
3.5 ml Glycerol (VWR)
35 µl 3 M Tris-HCl pH 8 (VWR)
ddH_2O was added to the final volume of 10ml.

Gibson Master Mix 2x (1.25ml)
50 µl NAD^+ (NEB)
50 µl 10mM dNTP (NEB)
25 µl 1 M DTT (filter sterilized) (Sigma-Aldrich)
405 µl Isothermal Start Mix (filter sterilized)
1 µl T5 exonuclease (NEB)
31.25 µl Phusion® High-Fidelity DNA polymerase (NEB)
250 µl Taq Ligase (NEB)
437.57 µl ddH_2O
Mixed gently by pipetting and stored as 100 µl aliquot at - 8^{0}C.

Glucose solution 40% (100ml)
100 g Glucose (VWR)
Dissolved in

IPTG 1M (10ml)
2.3829 g IPTG (G Biosciences)
10 ml ddH_2O was added to the final volume of 10ml.
Filtered through 0.42 µm filter, stored at -20^{0}C. Used at the final concentration of 0.1mM, 0.5mM and 1mM.

Isothermal Start Mix (3.15ml)
1.5 g PEG_{8000} (Sigma)
150 µl 2 M $MgCl_2$ (VWR)
3 ml 1 M Tris-HCL pH 8 (VWR)

PBS 10x (1L)
81.8 g NaCl (VWR)
14.2 g Na_2HPo_4 (VWR)
20.2 g KCl (VWR)
2. 45 g KH_2PO_4 (VWR)
The buffer autoclaved using P03 autoclave program for sterilizing the liquid.

TAE Buffer 50x (1L)
242 g Tris base (Sigma)
20.81 g EDTA (AppliChem)
57.1 ml Glacial acetic acid (Sigma-Aldrich)
The volume adjusted to 1L by adding ddH_2O.

2.Media used in this study

All the media used in this study, cited in material and methods

1. <u>Autoinduction media (studier)</u>

Autoinduction media used for induction of expression from the *lac* promoter in *E coli* BL21 Gold in biofilm assay and it is composed of the ingredient below:

<u>Zy media (1L)</u>
10 g Tryptone enzymatic digest from casein) (Sigma)
5 g Yeast extract (VWR)
950 ml ddH_2O was added
Autoclaved using P03 autoclave program for sterilizing the liquid.

<u>5052 50x(0.5L)</u>
25 g Glucose (VWR)
100 g Lactose (VWR)
0.5% Glycerol (VWR)

<u>$MgSO_4$ 1M (0.5L)</u>
123.24 g $MgSO_4$ (Merck)
Dissolved in 400 ml ddH_2O then the volume adjusted to 0.5L by adding ddH_2O.
Autoclaved using P03 autoclave program for sterilizing the liquid.

<u>NPS 20x (1L)</u>
66.07 g of 0.5 M $(NH_4)_2SO_4$ (Merck)
136.08 g 1M KH_2PO_4 (Merck)
177.99 g 1 M Na_2HPO_4 (VWR)
Dissolved in 900 ml ddH_2O then the volume adjusted to 1L by adding ddH_2O.
Autoclaved using P03 autoclave program for sterilizing the liquid.

#For making 200ml autoinduction media ZYP_5052 we need: -
5000 µl 1x NPS
2000 µl 1x 5052
100 µl (25 µg/ml) chloramphenicol
100 µl 1 mM $MgSO_4$
92.8 ml Zy media

2. <u>PA-0.5G defined media</u>

In order to get rid of washing step with 1xPBS, clear media is used for sedimentation assay. Clear media is composed of ingredients below: -

<u>NPS 20x (1L)</u>
66.07 g of 0.5 M $(NH_4)_2SO_4$ (Merck)
136.08 g 1M KH_2PO_4 (Merck)
177.99 g 1 M Na_2HPO_4 (VWR)
Dissolved in 900 ml ddH_2O then the volume adjusted to 1L by adding ddH_2O.
Autoclaved using P03 autoclave program for sterilizing the liquid.

40% Glucose (100ml)
40 g Glucose (VWR)
Dissolved in 80 ml ddH_2O then the volume adjusted to 100ml by adding ddH_2O
Filtered through 0.42 µm filter, stored at RT.

$MgSO_4$ 1M (0.5L)
123.24 g $MgSO_4$ (Merck)
Dissolved in 400 ml ddH_2O then the volume adjusted to 0.5L by adding ddH_2O.
Autoclaved using P03 autoclave program for sterilizing the liquid.

Amino acids 18aa(90ml)
1 g of each amino acids listed below were dissolved in 90ml ddH_2O, for o/n in the cold room. Next day the solution filtered through 0.42 µm filter and stored 15ml as aliquot at -20^0C. The amino acids used were: -
Na Glu, Asp, Lys-HCl, Arg-HCl, His-HCl, Ala, Pro, Gly, Thr, Ser, Gln, Asn, Val, Leu, Ile, Phe, Trp, and methionine.

20 000 x Micronutrients
Ammonium molybdate 6×10^{-5} M
Boric acid 8×10^{-3} M
Cobalt chloride 6×10^{-4} M
Copper sulphate 2×10^{-4} M
Manganese chloride 1.6×10^{-3} M
Zinc chloride 2×10^{-4} M
Autoclaved using P03 autoclave program for sterilizing the liquid.

For making 200ml PA-0.5G defined media we need the following below:
184.5 ml ddH_2O
200 µl 1 M $MgSO_4$ (Merck)
1 µl 20 000 x Micronutrients
2.5 ml 40% Glucose (VWR)
10 ml 20x NPS
10 µl 5 µg/ml Thiamin
2 ml 18 amino acid
200 µl 25 µg/ml chloramphenicol

3. LB media (1L)
5 g yeast extract (VWR)
10 g Tryptone (Sigma-Aldrich)
10 g NaCl (VWR)
1L ddH_2O was addedAutoclaved using P03 autoclave program for sterilizing the liquid.

4. SOC media (1L)
5 g yeast extract (VWR)
20 g Tryptone (Sigma-Aldrich)
0.5 g NaCl(VWR)
20 mM Glucose (VWR)
Autoclaved using P03 autoclave program for sterilizing the liquid.